

GenCore version 5.1.6.
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 23:43:34 ; Search time 36 Seconds

(without alignments)
958.564 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIIQAHLSLCLMLYLATG.....DKGDNVESQAASVVKPLRS 832

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt.40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4321	100.0	832	CADH_HUMAN	Q12864 homo sapien
2	3450	79.8	827	CADH_MOUSE	Q97100 mus musculu
3	3414	79.0	827	CADH_RAT	P55281 rattus norv
4	954	22.1	829	CADG_RABIT	Q28634 oryctolagus
5	922.5	21.3	829	CADG_HUMAN	Q75309 homo sapien
6	877.5	20.3	830	CADG_MOUSE	Q88338 mus musculu
7	715	16.5	913	CAD4_CHICK	P24503 gallus gall
8	705	16.3	913	CAD4_MOUSE	P39038 mus musculu
9	699.5	16.2	906	CAD2_HUMAN	P19022 homo sapien
10	696	16.1	912	CAD2_CHICK	P10288 gallus gall
11	694.5	16.1	906	CAD2_MOUSE	P15116 mus musculu
12	688.5	15.9	877	CAD2_BOVIN	P19534 bos taurus
13	686.5	15.9	906	CAD2_RAT	Q92193 rattus norv
14	684.5	15.8	916	CAD4_HUMAN	P55283 homo sapien
15	674.5	15.6	902	DSC3_MOUSE	P55292 mus musculu
16	671.5	15.5	896	DSC3_HUMAN	Q14574 homo sapien
17	667.5	15.4	904	DSC2_HUMAN	Q02487 homo sapien
18	663.5	15.4	906	CAD0_XENLA	P33147 xenopus lae
19	661	15.3	896	DSC3_BOVIN	Q28060 bos taurus
20	657	15.2	893	CADH_HUMAN	P55290 homo sapien
21	653.5	15.1	863	DSC2_BOVIN	P33545 bos taurus
22	647.5	15.0	905	CADN_XENLA	P20310 xenopus lae
23	647	15.0	714	CADH_MOUSE	Q9WTC5 mus musculu
24	642	14.9	895	DSC3_MOUSE	P55860 mus musculu
25	634	14.7	887	CAD1_CHICK	P08641 gallus gall
26	630.5	14.6	712	CADH_CHICK	P33150 gallus gall
27	609	14.1	886	CAD1_RAT	Q97074 rattus norv
28	608	14.1	790	CADI_HUMAN	Q13634 homo sapien
29	605	14.0	4590	FATH_HUMAN	Q14517 homo sapien
30	604	14.0	884	CADI_MOUSE	P09803 mus musculu
31	601.5	13.9	3579	STAN_DRONE	Q9V588 drosophila
32	597.5	13.8	732	CADL_CHICK	P33145 gallus gall
33	595.5	13.8	886	DSC1_MOUSE	P55849 mus musculu

34	591	13.7	882	CAD1_HUMAN	P12830 homo sapien
35	590	13.7	872	CAD1_XENLA	P30944 xenopus lae
36	590	13.6	784	CAD1_MOUSE	P33146 mus musculu
37	585.5	13.6	893	DSC1_BOVIN	Q01107 bos taurus
38	580.5	13.4	880	CAD1_XENLA	P33148 xenopus lae
39	576	13.3	880	CAD1_HUMAN	P33148 xenopus lae
40	574.5	13.3	796	CADH_MOUSE	Q9WTC5 mus musculu
41	574.5	13.3	894	DSC1_HUMAN	Q08554 homo sapien
42	571.5	13.2	789	CADH_CHICK	P79995 gallus gall
43	571	13.2	3034	CLRI_MOUSE	Q35161 mus musculu
44	567.5	13.1	788	CADH_HUMAN	Q9Y618 homo sapien
45	565	13.1	814	CADH_HUMAN	P55291 homo sapien

ALIGNMENTS

```

RESULT 1
ID      CADH_HUMAN      STANDARD;      PRT;      832 AA.
AC      Q12864; Q15336;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DE      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin)
DE      (Intestinal peptide-associated transporter HPT-1).
GN      CDH17.
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC      TISSUE=Colon adenocarcinoma;
RX      MEDLINE=94204643; PubMed=8153632;
RA      Dantzig A.H., Hoskins J., Tabae L.B., Bright S., Shepard R.L.,
RA      Jenkins I.L., Duckworth D.C., Sportsman J.R., Mackensen D.,
RA      Rostek P.R. Jr., Skatrud P.L.;
RT      "Association of intestinal peptide transport with a protein related to
RT      the cadherin superfamily.";
RL      Science 264:430-433 (1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Boettlinger A., Kreft B., Fieger C., Dlouhy B., Berndorf D.,
RA      Goessens R., Tauber P.;
RT      "Molecular cloning of human LI-cadherin: evidence for a novel type
RT      of cadherin within the cadherin superfamily.";
RL      Submitted (DEC-1994) to the EMBL/GenBank/DBD databases.
CC      -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC      THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC      MANNER IN CONNECTING CELLS. CADHERINS MAY THUS CONTRIBUTE TO THE
CC      SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
CC      IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
CC      INVOLVED IN INTESTINAL PEPTIDE TRANSPORT.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC      -1- TISSUE SPECIFICITY: EXPRESSED IN THE GASTROINTESTINAL TRACT AND
CC      PANCREATIC DUCT. NOT DETECTED IN KIDNEY, LUNG, LIVER, BRAIN,
CC      ADRENAL GLAND AND SKIN.
CC      -1- SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U07969; AAA19021.1; -
DR      EMBL; X83228; CAA58231.1; -
DR      HSSP; P15116; INCO.
DR      Genew; HGNC:1756; CDH17.
DR      MIM; 603017; -.

```

InterPro: IPR002126; Cadherin.
 DR Pfam; PF000028; Cadherin; 7.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 3.
 DR PROSITE; PS0268; CADHERIN 2; 6.
 KM Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 832 CADHERIN-17.
 FT DOMAIN 23 787 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 788 808 POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 809 832 POTENTIAL.
 FT DOMAIN 129 128 CADHERIN 1.
 FT DOMAIN 129 244 CADHERIN 2.
 FT DOMAIN 245 340 CADHERIN 3.
 FT DOMAIN 341 449 CADHERIN 4.
 FT DOMAIN 450 566 CADHERIN 5.
 FT DOMAIN 567 667 CADHERIN 6.
 FT DOMAIN 668 777 CADHERIN 7.
 FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 722 722 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 115 115 E -> K (IN REF. 2).
 FT CONFLICT 446 446 I -> T (IN REF. 2).
 FT CONFLICT 734 734 D -> E (IN REF. 2).
 FT CONFLICT 739 739 A -> E (IN REF. 2).
 SQ SEQUENCE 832 AA; 92147 MW; 66860608BA1277 CAC64;
 Query Match 100.0%; Score 4321; DB 1; Length 832;
 Best Local Similarity 100.0%; Pred. No. 7.1e-256;
 Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

541 FGKYNASSFAKFTLLVTPVNEAPQPSQHFQAKVEDVAIGTKGVNTAKDEGLDISY 600
 541 FGKYNASSFAKFTLLVTPVNEAPQPSQHFQAKVEDVAIGTKGVNTAKDEGLDISY 600
 601 SLRGDTGWLKIDHTVGEIPSAVPLDRBAGSPYRVQVATVEGSSLSVSEFHLIMDV 660
 601 SLRGDTGWLKIDHTVGEIPSAVPLDRBAGSPYRVQVATVEGSSLSVSEFHLIMDV 660
 601 SLRGDTGWLKIDHTVGEIPSAVPLDRBAGSPYRVQVATVEGSSLSVSEFHLIMDV 660
 661 NDNPRLAKDYTGIFCHPLSAGSLIFEXTDDOHLFRGPHFTSLGSSGLNDMEVSK 720
 661 NDNPRLAKDYTGIFCHPLSAGSLIFEXTDDOHLFRGPHFTSLGSSGLNDMEVSK 720
 721 INGTHTLSTRHTDPEERAVVVLIRINDGSRPLGIVSLPTFCSCVEGSCFRPAHQDT 780
 721 INGTHTLSTRHTDPEERAVVVLIRINDGSRPLGIVSLPTFCSCVEGSCFRPAHQDT 780
 781 GIFTVGAVGILLTTLVIGIILAVVFIKDKXGKDNVESAQASEVKPLRS 832
 781 GIFTVGAVGILLTTLVIGIILAVVFIKDKXGKDNVESAQASEVKPLRS 832
 781 GIFTVGAVGILLTTLVIGIILAVVFIKDKXGKDNVESAQASEVKPLRS 832
 RESULT 2
 CADD MOUSE STANDARD; PRT; 827 AA.
 ID CADD MOUSE
 AC 09R100;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin) (BILL-cadherin) (P130).
 GN CHD17.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Intestine;
 RA Angres B., Kim L., Tauber R.;
 RT "LI-cadherin gene expression during intestinal development.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.; SEQUENCE OF 26-33; 52-58; 74-81; 117-123 AND
 RP 490-509, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC STRAIN=BALB/c; TISSUE=Rectal liver;
 RX MEDLINE=20469471; PubMed=10906147;
 RA Ohnishi K., Shimizu T., Kanaeayama H., Melchers F.;
 RT "The identification of a nonclassical cadherin expressed during B cell
 development and its interaction with surrogate light chain.";
 RL J. Biol. Chem. 275:31134-31144(2000).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES. IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
 CC IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IS FOUND IN INTESTINE WITH
 CC LOWER EXPRESSION IN SPLEEN, BONE MARROW, LUNG AND TESTIS. NO
 CC EXPRESSION DETECTED IN LIVER, KIDNEY, HEART, BRAIN OR SKELETAL
 CC MUSCLE. EXPRESSED IN PRECURSOR B-CELLS AND MYELOID CELLS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION INCREASES IN PRO- AND PRE-B-I
 CC CELLS, DECREASES IN LARGE AND SMALL PRE-B-II CELLS, AND INCREASES
 CC AGAIN IN IMMATURE AND MATURE B-CELLS.
 CC -1- SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).

CC -----
 DR EMBL; AF177669; AAD51125.1; -
 DR EMBL; D87912; BAB03264.1; -
 DR HSSP; P15116; INCI.
 DR MGD; MG1:1095414; Cdh17.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 7.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS50268; CADHERIN_2; 5.
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal; Transport.
 FT SIGNAL 1 25
 FT CHAIN 26 827
 FT DOMAIN 26 786
 FT TRANSMEM 787 807
 FT DOMAIN 808 827
 FT DOMAIN 29 127
 FT DOMAIN 128 243
 FT DOMAIN 244 339
 FT DOMAIN 340 448
 FT DOMAIN 449 565
 FT DOMAIN 566 666
 FT DOMAIN 667 776
 FT CARBOHYD 148 148
 FT CARBOHYD 249 249
 FT CARBOHYD 418 418
 FT CARBOHYD 545 545
 FT CARBOHYD 573 573
 FT CARBOHYD 586 586
 FT CARBOHYD 721 721
 SQ SEQUENCE 827 AA; 91645 MW; CDCECEAS7652B58 CRC64;

Query Match 79.8%; Score 3450; DB 1; Length 827;
 Best Local Similarity 79.2%; Pred. No. 8.5e-203;
 Matches 658; Conservative 75; Mismatches 94; Indels 4; Gaps 1;

2 ILQAHLSCLMLYATAGGEGKFGPLKPMPTFSYEGOEPSQIFQKAMPATPE 61
 1 MVAHQHFLCLILYLCGIGEBKSGPLKPMPTFSIFEOEPSQVFPKTPPAVTFE 60
 62 LTGETDNIFVIEREGLLYNNALDRSTRTHNLQVALANGIIVGSPVITTEVKDIND 121
 61 LTGETDGIKIEKOGLLYHTRALDRTRAVHHQLALDSHGAIVGSPVITTEVKDIND 120
 122 NRPTFLOSKKEGVRQNSRPGKPFLLYNATLDDPATNGQLYYQIYVQLPMINWYFQ 181
 121 NRPTFLOSKKEGVRQNSRPGKPFMYNATLDDPATNGQLFYQIYVQLPQINDWYFQ 180
 182 INKTGASLTREGSOELNPAKNPSYLVTSYKDMGQSGNSPSDTSVDIITENIWK 241
 181 IDKTGASLTREGSOELNPAKNPSYLVTSYKDMGQSGNSPSDTSVDIITENIWK 240
 242 PKEVENENSTDPHPKITQVRNMWDGAGYSLVDKELPRFPSSIDQEGDIYVTOPLDRE 301
 241 PEVEIRENSTDPHPKITQVQWMDPGAGYSLVNDKELSPFPSSIDQEGDIYVTOPLDRE 300
 302 EKDAYFYAVAKRGYKPLSPLEIHYKVDINDNPTCPSPYTVFEVQENRIGNSIGT 361
 301 EKNSHFFPFAKDKNGKPLAYPEIYKVIDINDNPTCPSPYTVFEVQENRIGNSIGT 360
 362 LTAHDEDEENTANSFLNRYIVEQTPKLPMQGLFLIQTYGMLQARQSLLKQOTPOYNT 421
 361 FEAHDEDEANINISILKYKLVDDTPKVPDGLFLIGEBYKSVLSQSLKQSPQYNIS 420
 422 IEVSDKDFKLACVQAINVIDINDQIPIFEKSDYGNLTAEADNIGSTIITQATDADEPF 481
 421 IEVSDVDFKLACVQAINVIDINDQIPIFEKSDYGNLTAEADNIGSTIITQATDADEPF 480
 482 TGSSKLLIHLIKDSEGRGLGVDDPTHTNGYVILKKPLDFETIAVSNYFKAEPEPLVF 541
 481 TGSSKLLIYKLVQDTEGRLEVVDDPTHTNGYVILKKPLDFETQVSSIVFOAENPEPLVF 540

QY 542 GVKKNASPAKFLIYTDVNEAPQSOHVQAKYSEDAIGTKYGVNTAKDPGLDYSYS 601
 DB 541 GIEYNASSPASFELIYTDVNEAVPVFPQIRQANVSEDAVAGSVGNTADPGLDYSYS 600
 QY 602 LRGDTRGWLKIDYVTEIESVAPLDEBAGSPYRVQVATEVGGSSLSYSEPHILMDVN 661
 DB 601 LKGNMRGWLKIDYVTEIESVAPLDEBAGSPYRVQVATEVGGSSLSSTADPHLVITDVN 660
 QY 662 DNPRLAKYTGGLFCHPLSAPGSLPEATDDQHLFRGHFPFSLGSGSLQNDWEYSKI 721
 DB 661 DNPRLAKYTGGLFCHPLSAPGSLPEATDDQHLFRGHFPFSLGSGSLQNDWEYSKI 720
 QY 722 NGTHARLSTHTDPEERAYVVLIRINDGGRPLEGIVSLPVTFCSCVBSGCFRPAHQNG 781
 DB 721 NGTHARLSTHTDPEERAYVVLIRINDGGRPLEGIVSLPVTFCSCVBSGCFRPAHQNG 780
 QY 782 IPTVGNAGVILLTLLVIGIILAVFIRIKDKGKNVESAQAQSEVKPLRS 832
 DB 781 IPTVGNAGVILLTLLVIGIILAVFIRIKDKGKNVESAQAQSEVKPLRS 827

RESULT 3
 ID CADH_RAT STANDARD; PRT; 827 AA.
 AC P55281;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin).
 GN CDH17.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=9426966; PubMed=8207063;
 RA Berndorff D., Gessner R., Kreft B., Schnoy N., Lajous-Better A.-M.,
 RA Loch N., Reutter W., Hortsch M., Tauber R.;
 RT "Liver-intestine cadherin: molecular cloning and characterization of
 RT a novel Ca(2+)-dependent cell adhesion molecule expressed in liver
 RT and intestine";
 RL J. Cell Biol. 125:1353-1369(1994).
 CC - FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. LI-CADHERIN MAY HAVE A ROLE
 CC IN THE MORPHOLOGICAL ORGANIZATION OF LIVER AND INTESTINE.
 CC - SUBCELLULAR LOCATION: TYPE I membrane protein.
 CC - TISSUE SPECIFICITY: LIVER AND INTESTINE.
 CC - SIMILARITY: CONTAINS 7 CADHERIN DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X78997; CAA5631.1; -
 DR HSSP; P15116; INCI.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 7.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 6.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS50268; CADHERIN_2; 6.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL 1 21
 FT SIGNAL POTENTIAL.

```

FT CHAIN 22 827 CADERIN-17.
FT DOMAIN 22 786 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 787 807 POTENTIAL.
FT DOMAIN 808 827 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 29 127 CADERIN 1.
FT DOMAIN 128 243 CADERIN 2.
FT DOMAIN 244 339 CADERIN 3.
FT DOMAIN 340 448 CADERIN 4.
FT DOMAIN 449 565 CADERIN 5.
FT DOMAIN 566 666 CADERIN 6.
FT DOMAIN 667 776 CADERIN 7.
FT CARBOHYD 148 148 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 249 249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 827 AA; 91862 MW; E1A0B03646562C01 CRC64;

Query Match 79.0%; Score 3414; DB 1; Length 827;
Best Local Similarity 78.8%; Pred. No. 1,3e-200;
Matches 655; Conservative 70; Mismatches 102; Indels 4; Gaps 1;

2 IIAQHSLCLIMLYATGYGQEGKPSGLKPMTEFSIYEGQEPGQIIFQKAPNPAVTFE 61
1 MWSAQHFLCLLYLGAVGQEGKPSGLKPMTEFSIYEGQEPGQIIFQKAPNPAVTFE 60
62 LTGETDNI FVIEREGLLYNNALDRSTRNLQVALDANGIIVEGVPITIEVKIND 121
61 LTGETDGI FKIEKGGLYHTRVLDRETRAVAHLLQALMDSGALVDGPVPIIEVKIND 120
122 NRPFLLQSKIEGSRONSRPEKPFLLYNATLDLDPATPNQGLYQIYIQLPMINWYFQ 181
121 NRPFLLQSKIEGSRONSRPEKPFLLYNATLDLDPATPNQGLYQIYIQLPMINWYFQ 180
182 INNTGASLITREGSOELNPAKNPSYMLVISVKGMGQSGSFSDTTSVDITVENIWK 241
181 INNTGASLITREGSOELNPAKNPSYMLVISVKGMGQSGSFSDTTSVDITVENIWK 240
242 PKPEVWENSTDPHPIKITQVRANDPGQVSLVDKEKLPRPFISIDEGDIYVQPLDRE 301
241 PEPEIHEHNLDPPIKITQVRANDPGQVSLVDKEKLPRPFISIDEGDIYVQPLDRE 300
302 EKDAVYVYAAKDEYKPLSPLEIHYKVKDINNPPICSPVYVFEQENRERKGNISGT 361
301 EKSHVFPATKXDKNGKPLAYPLBIHYKVIDINNPPICSLQVTVFEQENRERKGNISGT 360
362 LTAHDBRENTANSFLNRYIVEQTPKLPMDGLFIQVAGKLOAKOSLKKQDTPQYNLT 421
361 PAHDMDBANNINSILKRLVDQTPKPSDBLFLIDEGKGVQGLKSKQDSPQYNLT 420
422 IEVSDKDEKTLCFVQVINVIDINDQIPFEKSDYGNLTAEEDTNGISTILLTQATDADPE 481
421 VEVSDDIFKTLCSQVNVVIDINDQIPFERSDYGSKTLSEDTAIGSTILLTQATDADPE 480
482 TGSSKILYHIKGSSEGLGYDTPHNTYGVIIKKPLDERTAAVSNVFAENPEPIYF 541
481 TGSSKILYHIKGSSEGLGYDTPHNTYGVIIKKPLDERTAAVSNVFAENPEPIYF 540
542 GVKYNASFAKFTLVTDVNEAPFSQHVFOAKYSEDAVATKYGAVTAKPREGDLSYS 601
541 GIEYNASSFASFELTVTDVNEAPFSQHVFOAKYSEDAVATKYGAVTAKPREGDLSYS 600
602 LRGDTRGMLKIDHYGEIYFVAPLDREAGSPYRQVAVATEVGGSSLSVSEFHLMDVN 661
601 LKDKRKGMLKIDHYGEIYFVAPLDREAGSPYRQVAVATEVGGSSLSVSEFHLMDVN 660
662 DNPRLAKDYGLGFPCPLSLAPGSLIFPATDDQGLFFGPHFTSLSGSGSLQNDWEYSKI 721
661 DNPRLAKDYGLGFPCPLSLAPGSLIFPATDDQGLFFGPHFTSLSGSGSLQNDWEYSKI 720
722 NGTHARLSTRHTDEERAYVVLIRINDGRPLBSGIVSLPVTFGSCVSGSCFRPAGHQTG 781

```

```

Db 721 NGTHARLSTRHTDEERAYVVLIRINDGRPLBSGIVSLPVTFGSCVSGSCFRPAGHQVG 780
Qy 782 IFTVGNAGVILTLTLLVIGIILAVTIRIKDKGKNVSAQSEVKPLRS 832
Db 781 IFTVGNAGVILTLTLLVIGIILAVTIRIKDKGKNVSAQSEVKPLRS 827

RESULT 4
CADD RABIT STANDARD; PRT; 829 AA.
ID CADD RABIT STANDARD; PRT; 829 AA.
AC Q28634;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Caderin-16 precursor (Kidney-specific caderin) (Ksp-caderin).
GN CDH16.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN (1)
RP SEQUENCE FROM N.A. AND OF SEQUENCE 570-586.
RC STRAIN=New Zealand white;
RX MEDLINE=95340560; Pubmed=7615566;
RA Thomson R.B., Igarashi P., Biemesderfer D., Kim R., Abu-Alfa A.,
RA Soleimani M., Atkinson P.S.;
RT "Isolation and cDNA cloning of Ksp-caderin, a novel kidney-specific
RT member of the caderin multigene family.";
RL J. Biol. Chem. 270:17594-17601(1995).
CC -1- FUNCTION: CADERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC -1- THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS. CADERINS MAY THIS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: KIDNEY-SPECIFIC. LIMITED TO THE BASOLATERAL
CC MEMBRANES OF RENAL TUBULAR EPITHELIAL CELLS.
CC -1- SIMILARITY: CONTAINS 6 CADERIN DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL; U28945; AAC48472.1; -.
DR HSSP; P09803; ISUH.
DR InterPro; IPR002126; Caderin.
DR Pfam; PF00028; caderin; 6.
DR SMART; SM00112; CA; 6.
DR PROSITE; PS00232; CADERIN_1; 2.
DR PROSITE; PS0268; CADERIN_2; 6.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 18
FT CHAIN 19 829
FT DOMAIN 19 786 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 787 807 POTENTIAL.
FT DOMAIN 808 829 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 126 CADERIN 1.
FT DOMAIN 131 235 CADERIN 2.
FT DOMAIN 242 336 CADERIN 3.
FT DOMAIN 341 449 CADERIN 4.
FT DOMAIN 445 564 CADERIN 5.
FT DOMAIN 569 665 CADERIN 6.
FT DOMAIN 666 786 CADERIN 7.
FT CARBOHYD 517 517 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 722 722 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 829 AA; 88827 MW; D2DF106C47A43B9 CRC64;

```


QY 297 PLDRBKDAVFAVAVAKDEYKPLSYPLHIVKVKDINDNPPCPSVTVFEVQENRLG 356
 DB 293 ELDRBAQAEYLLQVRAQNSHGBDYAAPLELHVLMWDENDVPIPPDPVTVPELSPPG 352
 QY 357 NSIGTLTAHNRDEBENTANSFLNRYIVEQTKLPMDG-LFLIQTAYAGMLQAKOSLKK-OD 414
 DB 353 TEVTRLSAEDADAGSPNSHVYVOLLSPEDVEGAFAQVDPPISSGVTLTGAVFLRAGON 412
 QY 415 TPQYNLTIEVSDKD-FKTLCTFOINVIDINDQIPREKSDYGNLTLAEDTNIGSTLTI 472
 DB 413 ILLIVLAMDLAGAAGCFSSICEVAVATDINDHAPEFTSQIGPISLPEDVEPGLVAML 472
 QY 473 QARDAD-EPTGSSKIL-YHIKGDSSRLGVDPDPHTNNGYVLIK-KPLDEFTAVSN 528
 DB 473 TALDADLEP--ARLMDPALERDTEGTGLDMEP--DSGHVRLCKLSTYBAASHE 527
 QY 529 IVFKAENPEPLFVGVKNASSFAFTLLIVDNEAPQPSQHPQAKYSEDAIGTKVNV 588
 DB 528 VVVVVGVAKTV-GPGPGPATATVTVLVERVMPPLKDQDSYASVPISAPNSFILTI 586
 QY 589 TAKDPGLDISYLRGDTRGMLKIDHVTGEIFSVAPLD-REAGSPRYOVATEVGGSSL 647
 DB 587 QPSDPIRTRLFSLVNDSEGLCTEKSSGEVHTAQSLQAGPQDPTVTVLEAQTDEPRL 646
 QY 648 SSVSEFHLIMDVNDNPRRLAKOYTGLFCHPLSAPSLIF-EATDDQHLFRGPHFTFS 706
 DB 647 SASAPLVTHLPKAPPAALTLTAPVPSQVLCYTRPDHGLTVSGSKDPLASGHGP-XSFT 705
 QY 707 LG-SGSLQNDWEVSKINGTARLSTRHTDFEERAVVLLIRINDGRPLEGIVSLPYTFC 765
 DB 706 LGRPPTQQRDRLQTLNGSHAVLTLALHWEPREHIIIPVVSHNAQ--WMQLLVRYVC 762
 QY 766 SC-VESGCFRPAHQGTGIPYGMVGLITLTVYIGIILAVFRIKKKKNVESAQA 824
 DB 763 RCNVEGQCMRKVGMKMPKLSAVGLVGLTVAIGLILFIHTWMSRKXDPDQPADS 822
 QY 825 SEVK 828
 DB 823 VPLK 826
 RESULT 6
 CADG_MOUSE
 ID CADG_MOUSE STANDARD, PRT: 830 AA.
 AC 088338; Q9JL25;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-16 precursor (Kidney-specific cadherin) (Kcp-cadherin).
 GN Cdh16.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 RX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=98389630; PubMed=9721215;
 RA Thomson R.B., Ward D.C., Quaggin S.E., Igarashi P., Muckler Z.E.,
 RA Argonson P.S.;
 RT "cDNA cloning and chromosomal localization of the human and mouse
 RT isoforms of kcp-cadherin."
 RL Genomics 51:445-451(1998).
 RN [2]
 RP SEQUENCE OF 1-25 FROM N.A.
 RC STRAIN=129/Sy;
 RA Whyte D.A., Thomson R.B., Nix S.L., Zanjan R., Li C., Karp S.L.,
 RA Argonson P.S., Igarashi P.;
 RT "Kcp-cadherin gene promoter. I. Sequence analysis and renal
 RT epithelial-cell-specific activity."
 RL Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.

CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES.
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC -I- TISSUE SPECIFICITY: KIDNEY-SPECIFIC.
 CC -I- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF016271; AAC34254.1; -
 DR EMBL; AF118228; AAF28836.1; -
 DR HSSP; P09803; 1SUH.
 DR MGI; MGI:106671; Cdh16.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 6.
 DR SMART; SM00112; CA; 6.
 DR PROSITE; PS00232; CADHERIN_1; 1.
 DR PROSITE; PS00268; CADHERIN_2; 5.
 KM Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT SIGNAL. 1 21
 FT CHAIN 22 830
 FT DOMAIN 22 788
 FT TRANSLEM 789 809
 FT DOMAIN 810 830
 FT DOMAIN 27 128
 FT DOMAIN 133 237
 FT DOMAIN 244 338
 FT DOMAIN 343 451
 FT DOMAIN 457 566
 FT DOMAIN 571 667
 FT DOMAIN 668 788
 FT DOMAIN 519 519
 FT CARBOHYD 604 604
 FT CARBOHYD 724 724
 SQ SEQUENCE 830 AA; 89859 MW; 5EB1F06B5B7C2E25 CRC64;
 Query Match 20.3%; Score 877.5; DB 1; Length 830;
 Best Local Similarity 30.5%; Pred No. 5, 1e-46;
 Matches 245; Conservative 141; Mismatches 353; Indels 63; Gaps 22;
 QY 62 LTGETD---NIFVIERE-GLLYNRALDRETSHTNLQVALLDANGIIVEGVPITIEV 116
 DB 59 LSGDSMTADQNTFAVVDTSGLFVATRTLDREKAERYQLQVTLSEDRILMGQVLTVHV 118
 QY 117 KDINDNPTFLQSKRYGSVQNSRPGKRPYVNAATLDDPATNGQLYVIVQLPIMNN 176
 DB 119 KDENDQVQPSQAIYRAQLSQGTRPGVPFLFLEASGDAGTANSDIRFHLISQSPQPL 178
 QY 177 VMYFOINNKGAISLRBGSQELNPAKNPSPYVVISVKMGQGSNSFSTTSVDIIVTE 236
 DB 179 PMFQDLDPHLGALALSSGSTSLDPALEETQLIVQVKMGDDPSGHAQAT-VEISIVE 237
 QY 237 NIWKAPEVEMVNSTDPHPIKITOVAMNDPGAQYSLVDEKULPRFPFSIDQSDIYVQ 296
 DB 238 NSNAPELEPHLANLKVYPHSLAQVHWSGADVHQL---ESQPPGFVDTEGMLHVTM 294
 QY 297 PLDRBKDAVFAVAVAKDEYKPLSYPLHIVKVKDINDNPPCPSVTVFEVQENRLG 356
 DB 295 ELDRBAQAEYLLQVRAQNSHGBDYAAPLELHVLMWDENDVAPCSPPDPVTNPELSPPG 354
 QY 357 NSIGTLTAHNRDEBENTANSFLNRYIVEQTKLPMDG-LFLIQTAYAGMLQ-----AKOS 409
 DB 355 TEVTRLSAEDADAGSPNSHVYVOLLSPEDVEGAENKAFLDPTSGSVTLTGAVFLRAGOS 414
 QY 410 LKQDTPQVYNLTIEV--SDRQFTLCFQVINVIDINDQIPREKSDYGNLTLAEDTNIGS 467

Db 415 ILIO-----VLAVDLASSEGSLSTCEVTWMTDNNHAEFINSGIQPVTLPEDEKGA 469
Qy 468 TILTIQATDAD-BEFTGSSKIL-YHIKGDSEGLAVDTDPHTNTGVYIKKPLDEFTAA 525
Db 470 LVATMTADADLEP---AFRLMDEGDEGDEGFDLSWEPDSOHVQLARKULSTYAAP 526
Qy 526 VSNIVFAENPEPELVFGVKNASSFAKFTLLIVDNEAFQPSQHVQAKVSEDAVIGTKV 565
Db 527 DHKVVVVVSNIEELV-GPGCPAPATATVTLIVERVAVPLKIDESYETSIPVSTPASGL 565
Qy 586 GNTAPDDEGLDISYLRGDRGLKIDHTGTGELFSAFPLD-REAGSPYQVVA--TEV 642
Db 586 LTTQPSDPMKRTKFLSLVNDSEGLCIKEVSGEHTAQSLQAGAPGPTVYVLEAODTK 645
Qy 643 GGSLSVSESEFHLI-----LMDVNDPPLAKDYTLFCHPLSAGSLFEATDDQH 696
Db 646 PGLSTATVIVIHPLKASPVALTLASGPR-----HICTRQDYGVVVSGVSEPD 697
Qy 697 LFRGPHFTSLG-SGLSDNWEVSKINGTARLSTRHTDEBRAYVVLIRI-NDGGRPL 754
Db 698 ANRNGPYSFALGPPTVQDRMLQPLNDSHAYLTLALHWEVPGVYVAVVHDTMWOL 757
Qy 755 EGVSLPYTFCSC-VESSCFRPAHQGTPTVCAVGLITLTLVIGILAVPIRL--- 810
Db 758 Q---KVVIVCRNVEGQCKRKGKMGKPTKLSAVGLLGTLLAIGFILVFTHLALA 813
Qy 811 KXDKGDNVESAQSEKPLRS 832
Db 814 RKOLD-----QPADSVPLKA 828

RESULT 7
CAD4 CHICK STANDARD; PRT; 913 AA.
ID AC P245Q3;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cadherin-4 precursor (Retinal-cadherin) (R-cadherin) (R-CAD).
GN CDH4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=91299341; PubMed=1712604;
RA Inuzuka H., Miyatani S., Takeichi M.;
RT "R-cadherin: a novel Ca(2+)-dependent cell-cell adhesion molecule
RT expressed in the retina.";
RT Neuron 7:69-79(1991).
RL

-1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. MAY PLAY AN IMPORTANT ROLE
CC IN RETINAL DEVELOPMENT.
CC
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC
CC -1- TISSUE SPECIFICITY: EMBRYONIC BRAIN AND NEURONAL RETINA.
CC
CC -1- DEVELOPMENTAL STAGE: DETECTED ONLY AFTER SOME DEGREE OF NEURONAL
CC DIFFERENTIATION HAS TAKEN PLACE AND PERSISTS AT LEAST UP TO THE
CC NEWLY HATCHED STAGE.
CC
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL, D14459; BAA03356.1; -.
DR PIR, JH0424; ICHCR.
DR HSSP, P15116; INCI.
DR InterPro, IPR002126; Cadherin.
DR InterPro, IPR000233; Cadherin_C_term.
DR Pfam, PF00028; cadherin; 5.
DR Pfam, PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN 1; 3.
DR PROSITE; PS0268; CADHERIN 2; 5.
DR PROSITE; PS0268; CADHERIN 3; 5.
DR PROSITE; PS0268; CADHERIN 4; 5.
DR PROSITE; PS0268; CADHERIN 5; 5.
DR PROSITE; PS0268; CADHERIN 6; 5.
DR PROSITE; PS0268; CADHERIN 7; 5.
DR PROSITE; PS0268; CADHERIN 8; 5.
DR PROSITE; PS0268; CADHERIN 9; 5.
DR PROSITE; PS0268; CADHERIN 10; 5.
DR PROSITE; PS0268; CADHERIN 11; 5.
DR PROSITE; PS0268; CADHERIN 12; 5.
DR PROSITE; PS0268; CADHERIN 13; 5.
DR PROSITE; PS0268; CADHERIN 14; 5.
DR PROSITE; PS0268; CADHERIN 15; 5.
DR PROSITE; PS0268; CADHERIN 16; 5.
DR PROSITE; PS0268; CADHERIN 17; 5.
DR PROSITE; PS0268; CADHERIN 18; 5.
DR PROSITE; PS0268; CADHERIN 19; 5.
DR PROSITE; PS0268; CADHERIN 20; 5.
DR PROSITE; PS0268; CADHERIN 21; 5.
DR PROSITE; PS0268; CADHERIN 22; 5.
DR PROSITE; PS0268; CADHERIN 23; 5.
DR PROSITE; PS0268; CADHERIN 24; 5.
DR PROSITE; PS0268; CADHERIN 25; 5.
DR PROSITE; PS0268; CADHERIN 26; 5.
DR PROSITE; PS0268; CADHERIN 27; 5.
DR PROSITE; PS0268; CADHERIN 28; 5.
DR PROSITE; PS0268; CADHERIN 29; 5.
DR PROSITE; PS0268; CADHERIN 30; 5.
DR PROSITE; PS0268; CADHERIN 31; 5.
DR PROSITE; PS0268; CADHERIN 32; 5.
DR PROSITE; PS0268; CADHERIN 33; 5.
DR PROSITE; PS0268; CADHERIN 34; 5.
DR PROSITE; PS0268; CADHERIN 35; 5.
DR PROSITE; PS0268; CADHERIN 36; 5.
DR PROSITE; PS0268; CADHERIN 37; 5.
DR PROSITE; PS0268; CADHERIN 38; 5.
DR PROSITE; PS0268; CADHERIN 39; 5.
DR PROSITE; PS0268; CADHERIN 40; 5.
DR PROSITE; PS0268; CADHERIN 41; 5.
DR PROSITE; PS0268; CADHERIN 42; 5.
DR PROSITE; PS0268; CADHERIN 43; 5.
DR PROSITE; PS0268; CADHERIN 44; 5.
DR PROSITE; PS0268; CADHERIN 45; 5.
DR PROSITE; PS0268; CADHERIN 46; 5.
DR PROSITE; PS0268; CADHERIN 47; 5.
DR PROSITE; PS0268; CADHERIN 48; 5.
DR PROSITE; PS0268; CADHERIN 49; 5.
DR PROSITE; PS0268; CADHERIN 50; 5.
DR PROSITE; PS0268; CADHERIN 51; 5.
DR PROSITE; PS0268; CADHERIN 52; 5.
DR PROSITE; PS0268; CADHERIN 53; 5.
DR PROSITE; PS0268; CADHERIN 54; 5.
DR PROSITE; PS0268; CADHERIN 55; 5.
DR PROSITE; PS0268; CADHERIN 56; 5.
DR PROSITE; PS0268; CADHERIN 57; 5.
DR PROSITE; PS0268; CADHERIN 58; 5.
DR PROSITE; PS0268; CADHERIN 59; 5.
DR PROSITE; PS0268; CADHERIN 60; 5.
DR PROSITE; PS0268; CADHERIN 61; 5.
DR PROSITE; PS0268; CADHERIN 62; 5.
DR PROSITE; PS0268; CADHERIN 63; 5.
DR PROSITE; PS0268; CADHERIN 64; 5.
DR PROSITE; PS0268; CADHERIN 65; 5.
DR PROSITE; PS0268; CADHERIN 66; 5.
DR PROSITE; PS0268; CADHERIN 67; 5.
DR PROSITE; PS0268; CADHERIN 68; 5.
DR PROSITE; PS0268; CADHERIN 69; 5.
DR PROSITE; PS0268; CADHERIN 70; 5.
DR PROSITE; PS0268; CADHERIN 71; 5.
DR PROSITE; PS0268; CADHERIN 72; 5.
DR PROSITE; PS0268; CADHERIN 73; 5.
DR PROSITE; PS0268; CADHERIN 74; 5.
DR PROSITE; PS0268; CADHERIN 75; 5.
DR PROSITE; PS0268; CADHERIN 76; 5.
DR PROSITE; PS0268; CADHERIN 77; 5.
DR PROSITE; PS0268; CADHERIN 78; 5.
DR PROSITE; PS0268; CADHERIN 79; 5.
DR PROSITE; PS0268; CADHERIN 80; 5.
DR PROSITE; PS0268; CADHERIN 81; 5.
DR PROSITE; PS0268; CADHERIN 82; 5.
DR PROSITE; PS0268; CADHERIN 83; 5.
DR PROSITE; PS0268; CADHERIN 84; 5.
DR PROSITE; PS0268; CADHERIN 85; 5.
DR PROSITE; PS0268; CADHERIN 86; 5.
DR PROSITE; PS0268; CADHERIN 87; 5.
DR PROSITE; PS0268; CADHERIN 88; 5.
DR PROSITE; PS0268; CADHERIN 89; 5.
DR PROSITE; PS0268; CADHERIN 90; 5.
DR PROSITE; PS0268; CADHERIN 91; 5.
DR PROSITE; PS0268; CADHERIN 92; 5.
DR PROSITE; PS0268; CADHERIN 93; 5.
DR PROSITE; PS0268; CADHERIN 94; 5.
DR PROSITE; PS0268; CADHERIN 95; 5.
DR PROSITE; PS0268; CADHERIN 96; 5.
DR PROSITE; PS0268; CADHERIN 97; 5.
DR PROSITE; PS0268; CADHERIN 98; 5.
DR PROSITE; PS0268; CADHERIN 99; 5.
DR PROSITE; PS0268; CADHERIN 100; 5.
DR PROSITE; PS0268; CADHERIN 101; 5.
DR PROSITE; PS0268; CADHERIN 102; 5.
DR PROSITE; PS0268; CADHERIN 103; 5.
DR PROSITE; PS0268; CADHERIN 104; 5.
DR PROSITE; PS0268; CADHERIN 105; 5.
DR PROSITE; PS0268; CADHERIN 106; 5.
DR PROSITE; PS0268; CADHERIN 107; 5.
DR PROSITE; PS0268; CADHERIN 108; 5.
DR PROSITE; PS0268; CADHERIN 109; 5.
DR PROSITE; PS0268; CADHERIN 110; 5.
DR PROSITE; PS0268; CADHERIN 111; 5.
DR PROSITE; PS0268; CADHERIN 112; 5.
DR PROSITE; PS0268; CADHERIN 113; 5.
DR PROSITE; PS0268; CADHERIN 114; 5.
DR PROSITE; PS0268; CADHERIN 115; 5.
DR PROSITE; PS0268; CADHERIN 116; 5.
DR PROSITE; PS0268; CADHERIN 117; 5.
DR PROSITE; PS0268; CADHERIN 118; 5.
DR PROSITE; PS0268; CADHERIN 119; 5.
DR PROSITE; PS0268; CADHERIN 120; 5.
DR PROSITE; PS0268; CADHERIN 121; 5.
DR PROSITE; PS0268; CADHERIN 122; 5.
DR PROSITE; PS0268; CADHERIN 123; 5.
DR PROSITE; PS0268; CADHERIN 124; 5.
DR PROSITE; PS0268; CADHERIN 125; 5.
DR PROSITE; PS0268; CADHERIN 126; 5.
DR PROSITE; PS0268; CADHERIN 127; 5.
DR PROSITE; PS0268; CADHERIN 128; 5.
DR PROSITE; PS0268; CADHERIN 129; 5.
DR PROSITE; PS0268; CADHERIN 130; 5.
DR PROSITE; PS0268; CADHERIN 131; 5.
DR PROSITE; PS0268; CADHERIN 132; 5.
DR PROSITE; PS0268; CADHERIN 133; 5.
DR PROSITE; PS0268; CADHERIN 134; 5.
DR PROSITE; PS0268; CADHERIN 135; 5.
DR PROSITE; PS0268; CADHERIN 136; 5.
DR PROSITE; PS0268; CADHERIN 137; 5.
DR PROSITE; PS0268; CADHERIN 138; 5.
DR PROSITE; PS0268; CADHERIN 139; 5.
DR PROSITE; PS0268; CADHERIN 140; 5.
DR PROSITE; PS0268; CADHERIN 141; 5.
DR PROSITE; PS0268; CADHERIN 142; 5.
DR PROSITE; PS0268; CADHERIN 143; 5.
DR PROSITE; PS0268; CADHERIN 144; 5.
DR PROSITE; PS0268; CADHERIN 145; 5.
DR PROSITE; PS0268; CADHERIN 146; 5.
DR PROSITE; PS0268; CADHERIN 147; 5.
DR PROSITE; PS0268; CADHERIN 148; 5.
DR PROSITE; PS0268; CADHERIN 149; 5.
DR PROSITE; PS0268; CADHERIN 150; 5.
DR PROSITE; PS0268; CADHERIN 151; 5.
DR PROSITE; PS0268; CADHERIN 152; 5.
DR PROSITE; PS0268; CADHERIN 153; 5.
DR PROSITE; PS0268; CADHERIN 154; 5.
DR PROSITE; PS0268; CADHERIN 155; 5.
DR PROSITE; PS0268; CADHERIN 156; 5.
DR PROSITE; PS0268; CADHERIN 157; 5.
DR PROSITE; PS0268; CADHERIN 158; 5.
DR PROSITE; PS0268; CADHERIN 159; 5.
DR PROSITE; PS0268; CADHERIN 160; 5.
DR PROSITE; PS0268; CADHERIN 161; 5.
DR PROSITE; PS0268; CADHERIN 162; 5.
DR PROSITE; PS0268; CADHERIN 163; 5.
DR PROSITE; PS0268; CADHERIN 164; 5.
DR PROSITE; PS0268; CADHERIN 165; 5.
DR PROSITE; PS0268; CADHERIN 166; 5.
DR PROSITE; PS0268; CADHERIN 167; 5.
DR PROSITE; PS0268; CADHERIN 168; 5.
DR PROSITE; PS0268; CADHERIN 169; 5.
DR PROSITE; PS0268; CADHERIN 170; 5.
DR PROSITE; PS0268; CADHERIN 171; 5.
DR PROSITE; PS0268; CADHERIN 172; 5.
DR PROSITE; PS0268; CADHERIN 173; 5.
DR PROSITE; PS0268; CADHERIN 174; 5.
DR PROSITE; PS0268; CADHERIN 175; 5.
DR PROSITE; PS0268; CADHERIN 176; 5.
DR PROSITE; PS0268; CADHERIN 177; 5.
DR PROSITE; PS0268; CADHERIN 178; 5.
DR PROSITE; PS0268; CADHERIN 179; 5.
DR PROSITE; PS0268; CADHERIN 180; 5.
DR PROSITE; PS0268; CADHERIN 181; 5.
DR PROSITE; PS0268; CADHERIN 182; 5.
DR PROSITE; PS0268; CADHERIN 183; 5.
DR PROSITE; PS0268; CADHERIN 184; 5.
DR PROSITE; PS0268; CADHERIN 185; 5.
DR PROSITE; PS0268; CADHERIN 186; 5.
DR PROSITE; PS0268; CADHERIN 187; 5.
DR PROSITE; PS0268; CADHERIN 188; 5.
DR PROSITE; PS0268; CADHERIN 189; 5.
DR PROSITE; PS0268; CADHERIN 190; 5.
DR PROSITE; PS0268; CADHERIN 191; 5.
DR PROSITE; PS0268; CADHERIN 192; 5.
DR PROSITE; PS0268; CADHERIN 193; 5.
DR PROSITE; PS0268; CADHERIN 194; 5.
DR PROSITE; PS0268; CADHERIN 195; 5.
DR PROSITE; PS0268; CADHERIN 196; 5.
DR PROSITE; PS0268; CADHERIN 197; 5.
DR PROSITE; PS0268; CADHERIN 198; 5.
DR PROSITE; PS0268; CADHERIN 199; 5.
DR PROSITE; PS0268; CADHERIN 200; 5.
DR PROSITE; PS0268; CADHERIN 201; 5.
DR PROSITE; PS0268; CADHERIN 202; 5.
DR PROSITE; PS0268; CADHERIN 203; 5.
DR PROSITE; PS0268; CADHERIN 204; 5.
DR PROSITE; PS0268; CADHERIN 205; 5.
DR PROSITE; PS0268; CADHERIN 206; 5.
DR PROSITE; PS0268; CADHERIN 207; 5.
DR PROSITE; PS0268; CADHERIN 208; 5.
DR PROSITE; PS0268; CADHERIN 209; 5.
DR PROSITE; PS0268; CADHERIN 210; 5.
DR PROSITE; PS0268; CADHERIN 211; 5.
DR PROSITE; PS0268; CADHERIN 212; 5.
DR PROSITE; PS0268; CADHERIN 213; 5.
DR PROSITE; PS0268; CADHERIN 214; 5.
DR PROSITE; PS0268; CADHERIN 215; 5.
DR PROSITE; PS0268; CADHERIN 216; 5.
DR PROSITE; PS0268; CADHERIN 217; 5.
DR PROSITE; PS0268; CADHERIN 218; 5.
DR PROSITE; PS0268; CADHERIN 219; 5.
DR PROSITE; PS0268; CADHERIN 220; 5.
DR PROSITE; PS0268; CADHERIN 221; 5.
DR PROSITE; PS0268; CADHERIN 222; 5.
DR PROSITE; PS0268; CADHERIN 223; 5.
DR PROSITE; PS0268; CADHERIN 224; 5.
DR PROSITE; PS0268; CADHERIN 225; 5.
DR PROSITE; PS0268; CADHERIN 226; 5.
DR PROSITE; PS0268; CADHERIN 227; 5.
DR PROSITE; PS0268; CADHERIN 228; 5.
DR PROSITE; PS0268; CADHERIN 229; 5.
DR PROSITE; PS0268; CADHERIN 230; 5.
DR PROSITE; PS0268; CADHERIN 231; 5.
DR PROSITE; PS0268; CADHERIN 232; 5.
DR PROSITE; PS0268; CADHERIN 233; 5.
DR PROSITE; PS0268; CADHERIN 234; 5.
DR PROSITE; PS0268; CADHERIN 235; 5.
DR PROSITE; PS0268; CADHERIN 236; 5.
DR PROSITE; PS0268; CADHERIN 237; 5.
DR PROSITE; PS0268; CADHERIN 238; 5.
DR PROSITE; PS0268; CADHERIN 239; 5.
DR PROSITE; PS0268; CADHERIN 240; 5.
DR PROSITE; PS0268; CADHERIN 241; 5.
DR PROSITE; PS0268; CADHERIN 242; 5.
DR PROSITE; PS0268; CADHERIN 243; 5.
DR PROSITE; PS0268; CADHERIN 244; 5.
DR PROSITE; PS0268; CADHERIN 245; 5.
DR PROSITE; PS0268; CADHERIN 246; 5.
DR PROSITE; PS0268; CADHERIN 247; 5.
DR PROSITE; PS0268; CADHERIN 248; 5.
DR PROSITE; PS0268; CADHERIN 249; 5.
DR PROSITE; PS0268; CADHERIN 250; 5.
DR PROSITE; PS0268; CADHERIN 251; 5.
DR PROSITE; PS0268; CADHERIN 252; 5.
DR PROSITE; PS0268; CADHERIN 253; 5.
DR PROSITE; PS0268; CADHERIN 254; 5.
DR PROSITE; PS0268; CADHERIN 255; 5.
DR PROSITE; PS0268; CADHERIN 256; 5.
DR PROSITE; PS0268; CADHERIN 257; 5.
DR PROSITE; PS0268; CADHERIN 258; 5.
DR PROSITE; PS0268; CADHERIN 259; 5.
DR PROSITE; PS0268; CADHERIN 260; 5.
DR PROSITE; PS0268; CADHERIN 261; 5.
DR PROSITE; PS0268; CADHERIN 262; 5.
DR PROSITE; PS0268; CADHERIN 263; 5.
DR PROSITE; PS0268; CADHERIN 264; 5.
DR PROSITE; PS0268; CADHERIN 265; 5.
DR PROSITE; PS0268; CADHERIN 266; 5.
DR PROSITE; PS0268; CADHERIN 267; 5.
DR PROSITE; PS0268; CADHERIN 268; 5.
DR PROSITE; PS0268; CADHERIN 269; 5.
DR PROSITE; PS0268; CADHERIN 270; 5.
DR PROSITE; PS0268; CADHERIN 271; 5.
DR PROSITE; PS0268; CADHERIN 272; 5.
DR PROSITE; PS0268; CADHERIN 273; 5.
DR PROSITE; PS0268; CADHERIN 274; 5.
DR PROSITE; PS0268; CADHERIN 275; 5.
DR PROSITE; PS0268; CADHERIN 276; 5.
DR PROSITE; PS0268; CADHERIN 277; 5.
DR PROSITE; PS0268; CADHERIN 278; 5.
DR PROSITE; PS0268; CADHERIN 279; 5.
DR PROSITE; PS0268; CADHERIN 280; 5.
DR PROSITE; PS0268; CADHERIN 281; 5.
DR PROSITE; PS0268; CADHERIN 282; 5.
DR PROSITE; PS0268; CADHERIN 283; 5.
DR PROSITE; PS0268; CADHERIN 284; 5.
DR PROSITE; PS0268; CADHERIN 285; 5.
DR PROSITE; PS0268; CADHERIN 286; 5.
DR PROSITE; PS0268; CADHERIN 287; 5.
DR PROSITE; PS0268; CADHERIN 288; 5.
DR PROSITE; PS0268; CADHERIN 289; 5.
DR PROSITE; PS0268; CADHERIN 290; 5.
DR PROSITE; PS0268; CADHERIN 291; 5.
DR PROSITE; PS0268; CADHERIN 292; 5.
DR PROSITE; PS0268; CADHERIN 293; 5.
DR PROSITE; PS0268; CADHERIN 294; 5.
DR PROSITE; PS0268; CADHERIN 295; 5.
DR PROSITE; PS0268; CADHERIN 296; 5.
DR PROSITE; PS0268; CADHERIN 297; 5.
DR PROSITE; PS0268; CADHERIN 298; 5.
DR PROSITE; PS0268; CADHERIN 299; 5.
DR PROSITE; PS0268; CADHERIN 300; 5.
DR PROSITE; PS0268; CADHERIN 301; 5.
DR PROSITE; PS0268; CADHERIN 302; 5.
DR PROSITE; PS0268; CADHERIN 303; 5.
DR PROSITE; PS0268; CADHERIN 304; 5.
DR PROSITE; PS0268; CADHERIN 305; 5.
DR PROSITE; PS0268; CADHERIN 306; 5.
DR PROSITE; PS0268; CADHERIN 307; 5.
DR PROSITE; PS0268; CADHERIN 308; 5.
DR PROSITE; PS0268; CADHERIN 309; 5.
DR PROSITE; PS0268; CADHERIN 310; 5.
DR PROSITE; PS0268; CADHERIN 311; 5.
DR PROSITE; PS0268; CADHERIN 312; 5.
DR PROSITE; PS0268; CADHERIN 313; 5.
DR PROSITE; PS0268; CADHERIN 314; 5.
DR PROSITE; PS0268; CADHERIN 315; 5.
DR PROSITE; PS0268; CADHERIN 316; 5.
DR PROSITE; PS0268; CADHERIN 317; 5.
DR PROSITE; PS0268; CADHERIN 318; 5.
DR PROSITE; PS0268; CADHERIN 319; 5.
DR PROSITE; PS0268; CADHERIN 320; 5.
DR PROSITE; PS0268; CADHERIN 321; 5.
DR PROSITE; PS0268; CADHERIN 322; 5.
DR PROSITE; PS0268; CADHERIN 323; 5.
DR PROSITE; PS0268; CADHERIN 324; 5.
DR PROSITE; PS0268; CADHERIN 325; 5.
DR PROSITE; PS0268; CADHERIN 326; 5.
DR PROSITE; PS0268; CADHERIN 327; 5.
DR PROSITE; PS0268; CADHERIN 328; 5.
DR PROSITE; PS0268; CADHERIN 329; 5.
DR PROSITE; PS0268; CADHERIN 330; 5.
DR PROSITE; PS0268; CADHERIN 331; 5.
DR PROSITE; PS0268; CADHERIN 332; 5.
DR PROSITE; PS0268; CADHERIN 333; 5.
DR PROSITE; PS0268; CADHERIN 334; 5.
DR PROSITE; PS0268; CADHERIN 335; 5.
DR PROSITE; PS0268; CADHERIN 336; 5.
DR PROSITE; PS0268; CADHERIN 337; 5.
DR PROSITE; PS0268; CADHERIN 338; 5.
DR PROSITE; PS0268; CADHERIN 339; 5.
DR PROSITE; PS0268; CADHERIN 340; 5.
DR PROSITE; PS0268; CADHERIN 341; 5.
DR PROSITE; PS0268; CADHERIN 342; 5.
DR PROSITE; PS0268; CADHERIN 343; 5.
DR PROSITE; PS0268; CADHERIN 344; 5.
DR PROSITE; PS0268; CADHERIN 345; 5.
DR PROSITE; PS0268; CADHERIN 346; 5.
DR PROSITE; PS0268; CADHERIN 347; 5.
DR PROSITE; PS0268; CADHERIN 348; 5.
DR PROSITE; PS0268; CADHERIN 349; 5.
DR PROSITE; PS0268; CADHERIN 350; 5.
DR PROSITE; PS0268; CADHERIN 351; 5.
DR PROSITE; PS0268; CADHERIN 352; 5.
DR PROSITE; PS0268; CADHERIN 353; 5.
DR PROSITE; PS0268; CADHERIN 354; 5.
DR PROSITE; PS0268; CADHERIN 355; 5.
DR PROSITE; PS0268; CADHERIN 356; 5.
DR PROSITE; PS0268; CADHERIN 357; 5.
DR PROSITE; PS0268; CADHERIN 358; 5.
DR PROSITE; PS0268; CADHERIN 359; 5.
DR PROSITE; PS0268; CADHERIN 360; 5.
DR PROSITE; PS0268; CADHERIN 361; 5.
DR PROSITE; PS0268; CADHERIN 362; 5.
DR PROSITE; PS0268; CADHERIN 363; 5.
DR PROSITE; PS0268; CADHERIN 364; 5.
DR PROSITE; PS0268; CADHERIN 365; 5.
DR PROSITE; PS0268; CADHERIN 366; 5.
DR PROSITE; PS0268; CADHERIN 367; 5.
DR PROSITE; PS0268; CADHERIN 368; 5.
DR PROSITE; PS0268; CADHERIN 369; 5.
DR PROSITE; PS0268; CADHERIN 370; 5.
DR PROSITE; PS0268; CADHERIN 371; 5.
DR PROSITE; PS0268; CADHERIN 372; 5.
DR PROSITE; PS0268; CADHERIN 373; 5.
DR PROSITE; PS0268; CADHERIN 374; 5.
DR PROSITE; PS0268; CADHERIN 375; 5.
DR PROSITE; PS0268; CADHERIN 376; 5.
DR PROSITE; PS0268; CADHERIN 377; 5.
DR PROSITE; PS0268; CADHERIN 378; 5.
DR PROSITE; PS0268; CADHERIN 379; 5.
DR PROSITE; PS0268; CADHERIN 380; 5.
DR PROSITE; PS0268; CADHERIN 381; 5.
DR PROSITE; PS0268; CADHERIN 382; 5.
DR PROSITE; PS0268; CADHERIN 383; 5.
DR PROSITE; PS0268; CADHERIN 384; 5.
DR PROSITE; PS0268; CADHERIN 385; 5.
DR PROSITE; PS0268; CADHERIN 386; 5.
DR PROSITE; PS0268; CADHERIN 387; 5.
DR PROSITE; PS0268; CADHERIN 388; 5.
DR PROSITE; PS0268; CADHERIN 389; 5.
DR PROSITE; PS0268; CADHERIN 390; 5.
DR PROSITE; PS0268; CADHERIN 391; 5.
DR PROSITE; PS0268; CADHERIN 392; 5.
DR PROSITE; PS0268; CADHERIN 393; 5.
DR PROSITE; PS0268; CADHERIN 394; 5.
DR PROSITE; PS0268; CADHERIN 395; 5.
DR PROSITE; PS0268; CADHERIN 396; 5.
DR PROSITE; PS0268; CADHERIN 397; 5.
DR PROSITE; PS0268; CADHERIN 398; 5.
DR PROSITE; PS0268; CADHERIN 399; 5.
DR PROSITE; PS0268; CADHERIN 400; 5.
DR PROSITE; PS0268; CADHERIN 401; 5.
DR PROSITE; PS0268; CADHERIN 402; 5.
DR PROSITE; PS0268; CADHERIN 403; 5.
DR PROSITE; PS0268; CADHERIN 404; 5.
DR PROSITE; PS0268; CADHERIN 405; 5.
DR PROSITE; PS0268; CADHERIN 406; 5.
DR PROSITE; PS0268; CADHERIN 407; 5.
DR PROSITE; PS0268; CADHERIN 408; 5.
DR PROSITE; PS0268; CADHERIN 409; 5.
DR PROSITE; PS0268; CADHERIN 410; 5.
DR PROSITE; PS0268; CADHERIN 411; 5.
DR PROSITE; PS0268; CADHERIN 412; 5.
DR PROSITE; PS0268; CADHERIN 413; 5.
DR PROSITE; PS0268; CADHERIN 414; 5.
DR PROSITE; PS0268; CADHERIN 415; 5.
DR PROSITE; PS0268; CADHERIN 416; 5.
DR PROSITE; PS0268; CADHERIN 417; 5.
DR PROSITE; PS0268; CADHERIN 418; 5.
DR PROSITE; PS0268; CADHERIN 419; 5.
DR PROSITE; PS0268; CADHERIN 420; 5.
DR PROSITE; PS0268; CADHERIN 421; 5.
DR PROSITE; PS0268; CADHERIN 422; 5.
DR PROSITE; PS0268; CADHERIN 423; 5.
DR PROSITE; PS0268; CADHERIN 424; 5.
DR PROSITE; PS0268; CADHERIN 425; 5.
DR PROSITE; PS0268; CADHERIN 426; 5.
DR PROSITE; PS0268; CADHERIN 427; 5.
DR PROSITE; PS0268; CADHERIN 428; 5.
DR PROSITE; PS0268; CADHERIN 429; 5.
DR PROSITE; PS0268; CADHERIN 430; 5.
DR PROSITE; PS0268; CADHERIN 431; 5.
DR PROSITE; PS0268; CADHERIN 432; 5.
DR PROSITE; PS0268; CADHERIN 433; 5.
DR PROSITE; PS0268; CADHERIN 434; 5.
DR PROSITE; PS0268; CADHERIN 435; 5.
DR PROSITE; PS0268; CADHERIN 436; 5.
DR PROSITE; PS0268; CADHERIN 437; 5.
DR PROSITE; PS0268; CADHERIN 438; 5.
DR PROSITE; PS0268; CADHERIN 439; 5.
DR PROSITE; PS0268; CADHERIN 440; 5.
DR PROSITE; PS0268; CADHERIN 441; 5.
DR PROSITE; PS0268; CADHERIN 442; 5.
DR PROSITE; PS0268; CADHERIN 443; 5.
DR PROSITE; PS0268; CADHERIN 444; 5.
DR PROSITE; PS0268; CADHERIN 445; 5.
DR PROSITE; PS0268; CADHERIN 446; 5.
DR PROSITE; PS0268; CADHERIN 447; 5.
DR PROSITE; PS0268; CADHERIN 448; 5.
DR PROSITE; PS0268; CADHERIN 449; 5.
DR PROSITE; PS0268; CADHERIN 450; 5.
DR PROSITE; PS0268; CADHERIN 451; 5.
DR PROSITE; PS0268; CADHERIN 452; 5.
DR PROSITE; PS0268; CADHERIN 453; 5.
DR PROSITE; PS0268; CADHERIN 454; 5.
DR PROSITE; PS0268; CADHERIN 455; 5.
DR PROSITE; PS0268; CADHERIN 456; 5.
DR PROSITE; PS0268; CADHERIN 457; 5.
DR PROSITE; PS0268; CADHERIN 458; 5.
DR PROSITE; PS0268; CADHERIN 459; 5.
DR PROSITE; PS0268; CADHERIN 460; 5.
DR PROSITE; PS0268; CADHERIN 461; 5.
DR PROSITE; PS0268; CADHERIN 462; 5.
DR PROSITE; PS0268; CADHERIN 463; 5.
DR PROSITE; PS0268; CADHERIN 464; 5.
DR PROSITE; PS0268; CADHERIN 465; 5.
DR PROSITE; PS0268; CADHERIN 466; 5.
DR PROSITE; PS0268; CADHERIN 467; 5.
DR PROSITE; PS0268; CADHERIN 468; 5.
DR PROSITE; PS0268; CADHERIN 469; 5.
DR PROSITE; PS0268; CADHERIN 470; 5.
DR PROSITE; PS0268; CADHERIN 471; 5.

QY 604 GDRGWLKIDHVTGEISVAPLDRPRA-----GSPRVQVATVGGSSLSVSEPHLLMD 659
 DB 545 SDBANMNTINATNGQITTAVALDRESYIKNNVTEATFELADNGIPASGTGLQIYLID 604
 QY 660 VDNPPRLADYDGLFCFCHPLSAGSLIFEAT-----DDOHLFRGPHFTSLGS--GSLQ 713
 DB 605 INDNAPLLEKRAOI-----CEKNLANVINTAADLIDRNV--GP-FVLELSPVPSAVR 656
 QY 714 NDWEVSKINGTHARLSTRHTDPEERAYVVLIRINDGRPPLEGIVLSVTFSCVSGSCF 773
 DB 657 KNNWITRLNGDYAQLSRIMVLEAGYVDPLIVTDSGNPPLYNTSIKVKPCDEGDC 716
 QY 774 RPAGHOTGIPVGAAGVILTLTLVIGIILAVFI 808
 DB 717 TTIG---AVAAAGTGAIITAILIITLITLWLL 748
 RESULT 8
 ID CAD4_MOUSE STANDARD; PRT; 913 AA.
 AC P39038;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-4 precursor (Retinal-cadherin) (R-cadherin) (R-CAD).
 GN C0H4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreatic islets;
 RX MEDLINE=94067164; PubMed=8247017;
 RA Hutton J.C., Christofori G., Chi W.Y., Edman U., Guest P.C.,
 RA Hanahan D., Kelly R.B.;
 RT "Molecular cloning of mouse pancreatic islet R-cadherin: differential
 RT expression in endocrine and exocrine tissue.";
 RT Mol. Endocrinol. 7:1151-1160 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94095672; PubMed=8270638;
 RA Matsunami H., Miyatani S., Inoue T., Copeland N.G., Gilbert D.,
 RA Jenkins N.A., Takeichi M.;
 RT "Cell binding specificity of mouse R-cadherin and chromosomal mapping
 RT of the gene.";
 RT J. Cell Sci. 106:401-409 (1993).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. MAY PLAY AN IMPORTANT ROLE
 CC IN RETINAL DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: DISTRIBUTED WIDELY IN MOUSE TISSUES WITH HIGH
 CC LEVELS PRESENT IN BRAIN, SKELETAL MUSCLE, AND THYMUS.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X69966; CAA49589.1; -
 DR EMBL; D14888; BAA03605.1; -
 DR PIR; A47543; A47543.
 DR HGSP; P15116; INCI.
 DR MGD; MGI:99218; Cdh4.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C_term.

DR Pfam; PF00028; cadherin; 5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN_1; 3.
 DR PROSITE; PS00268; CADHERIN_2; 5.
 KM Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KW Signal.
 FT 1 20
 FT PROPEP 21 166
 FT CHAIN 167 913
 FT DOMAIN 167 731
 FT TRANSMEM 732 753
 FT DOMAIN 754 913
 FT DOMAIN 167 274
 FT DOMAIN 275 389
 FT DOMAIN 390 504
 FT DOMAIN 505 610
 FT DOMAIN 611 721
 FT DOMAIN 721 885
 FT CARBOHYD 146 146
 FT CARBOHYD 280 280
 FT CARBOHYD 409 409
 FT CARBOHYD 554 554
 FT CARBOHYD 629 629
 FT CARBOHYD 658 658
 FT CARBOHYD 699 699
 SQ SEQUENCE 913 AA; 100030 MW; 1245A8CE8C38481 CRC64;
 Query Match 16.3%; Score 705; DB 1; Length 913;
 Best local Similarity 30.1%; Pred. No. 1.9e-35;
 Matches 178; Conservative 101; Mismatches 280; Indels 33; Gaps 12;
 QY 239 WKAPKVEVWENSTDPRIKITOVK---NDPBAQVSL--VDKXKLPRFFPSID-QEGDI 292
 DB 168 WVLP-PINVENENRGPFPOQLVIRSDKNDIPRYSITGVGADQPMVEVFNDSMGRM 226
 QY 293 YVQPLDREKDYVFAVAKDEYKPLSYELIHWKVDINDNPTCPSPVTVFEVOEN 352
 DB 227 YVTRPMDREBRASVHLRAHVDNKGKVENPDLIYIVDMNDRBEFINQVNGSDEG 286
 QY 353 ERLGNSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDGLFLIQYAGMLQAKOSLKK 412
 DB 287 SKPGTYVMTATADADSTTANGVWRVRYVTPQSPSQMFTINSETGDIIVVAAGLDR 346
 QY 413 QDTPQNLITREVDK-----FKTLCFVQINVIDINDQPIREKSDYGLTLAEDNIG 466
 DB 347 EKVQOVYVIVQATDMENLNYGLSNTATALITVTVANDNPEFTSTFAG--EVENRIE 404
 QY 467 STILTIQATADAPFTGSSKILYIIKGDSEGLGYDTPHTNGVYLIKPLDFETAAV 526
 DB 405 TVVANLTVMDRDQPHSMNNAVRIISGDSGFVSRTDVTNMEGAVTVKAYDELNRA 464
 QY 527 SNIVFAENPEPLVFGVKNVASSPAKFTLLVTVDNEAPQFSQHVQAKVSEDAVIGTVG 586
 DB 465 FMLTVMWSNQAPLASGIQMSFGSTAGVTISVTVDNEAPYSPSHKILRLREGVPAATAL 524
 QY 587 NVTAKDPEGL---DISLSLGDTRGWLKIDHVTGEISVAPLDRPRA-----GSPRVQVVA 639
 DB 525 TFSVDPDRMOQAVRYSKISDPANMVAHTNSGQITTAALDRESLYTKNNVTEATFLA 584
 QY 640 TVVGGSSLSVSEPHLLMDVNDPRLADYDGLFCFCHPLSAG--SLIFEATDDQHL 697
 DB 585 ADNGIPASGTGLQIYLIDINDAPLLEKRAOI-----CERPLNAINITAADADMDP 639
 QY 698 FRGPH-FTFSLGSGSLNDWEVSKINGTHARLSTRHTDPEERAYVVLIRINDGRPPLEG 756
 DB 640 NIGYVVELFPIPTTAKMTITRLNGDYAQLSRIMVLEAGYVDPLIVTDSGNPPLSN 699
 QY 757 IVSLPVTGSCVSGSCRPAGHOTGIPVGAAGVILTLTLVIGIILAVFI 808
 DB 700 TSVIKVAVCPDENGDCITVG---AVAAAGTGAIITAILIITLITLWLL 748


```

Db      203 ----DPPPTGIFINPISGOLSTYKPLDRELLRPHLRRAHYDINGNOVENPIDIYINYI 258
Qy      332 DINDNPCTCSPVTVPEVOENERLGNISIGTLTAHDREENTANSFLNYRIVECTPKLPM 391
Db      259 DMDNDNRPEFLHQWNSVPEGSRGTYVMVTVAIDADDPAALNGMLRYRLISQAFSTPSP 318
Qy      392 GLFLITVAGMLQAKOSLKKQOTPOYNLTIEVSDK-----FKLCEVQIVINDINQ 445
Db      319 NMFTINNETGDIITVAAGLDREKVOQYTLTIQATDMEGNFTYGLSTATAVITVTVDN 378
Qy      446 IPIF-EKSDYGNLTAEADTNGSTLTITQATDDEPTGSSKILYHIKDGSEGLGVP 504
Db      379 PPEFTANTFYGEV---PENRVDIVANLTVDKQDPHTPMANAYRISGDPYGRFALIT 435
Qy      505 DPHNTNGYVILIKKPLDFTAAVSNIVFKAENPEPLVFGVKYKNSAFKFTLITVDNEAP 564
Db      436 DPNNDGLTVVAVRIPDETNRMFVLVAENQVPLAKGIQHPPOSTATVSVTVDVNEP 495
Qy      565 QESQHVFOAKVSEDAVIGTVGNVTAADPEGL--DISYSLRQDTRGMLKIDHVTGEIPS 621
Db      496 YFAPNPKIRQEBGLHAGTWTLTITADDPDRYMOQNIRYKLSDBPANMLKIDEPVNGQITT 555
Qy      622 VAPLDREA---GSPYRVQVATEVGGSSLSVSEFHLIMDVNDNPPRLAKDYTGLEFC 677
Db      556 IAVLDHESPVYQNNYINATFLASDNGLIPMSGCTLOIYILDINDNAPOVL----- 606
Qy      678 HPLSA-----PGLIFBATDDOHLFRGPH-FTFSLGSGSLQNDWEVSKINGTHARLS 729
Db      607 -PEAEETCETPEPNSINIALDYDIDNAGPFAVDLPLSPVTLIKRMWTINRLNGDFQALN 665
Qy      730 TRHTDPEERAYVYLIRINDGRPPLEIVSLPTFCGC-VEGSCFRAGHGOTGTPV--- 785
Db      666 LKIKFLFAGIYEVIITITDSGNPKNSISILRVKVCQCSNQC-----TVYDRLVGA 718
Qy      786 GMAVGILTLVLVIGLIVAFIRIKKDKGKDNVESQASEVPR 829
Db      719 GLGTGATITALLCIIILILIVLVVWVKRDKEROKKOLLIDP 762

```

RESULT 12

```

CAD2_BOVIN
ID_CAD2_BOVIN STANDARD; PRT; 877 AA.
AC P19534;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neural-cadherin precursor (N-cadherin) (Cadherin-2) (Fragment).
GN CDH2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=30360979; PubMed=2390969;
RA Llaw C.W., Cannon C., Power M.D., Kiboneka P.K., Rubin L.L.;
RT "Identification and cloning of two species of cadherins in bovine
RT endothelial cells."
RL EMBO J. 9:2701-2708(1990).
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
CC NEURONAL RECOGNITION MECHANISM.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53615; CAA37677.1; -.
CC DR PIR; S11693; IUBOEN.
CC DR HSSP; P5116; INCI.
CC DR InterPro; IPR002126; Cadherin.
CC DR InterPro; IPR000233; Cadherin_C-term.
CC DR Pfam; PF00028; cadherin; 5.
CC DR Pfam; PF01049; cadherin_C-term; 1.
CC DR SMART; SM00112; CA; 5.
CC DR PROSITE; PS00232; CADHERIN_1; 3.
CC DR PROSITE; PS00268; CADHERIN_2; 5.
CC KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
CC FT NON TER 1 130
CC FT PROPEP <1 877
CC FT CHAIN 131 695
CC FT DOMAIN 131 695
CC FT TRANSMEM 696 717
CC FT DOMAIN 718 877
CC FT DOMAIN 131 238
CC FT DOMAIN 239 363
CC FT DOMAIN 364 468
CC FT DOMAIN 469 574
CC FT DOMAIN 575 685
CC FT DOMAIN 834 849
CC FT CARBOHYD 161 161
CC FT CARBOHYD 244 244
CC FT CARBOHYD 296 296
CC FT CARBOHYD 373 373
CC FT CARBOHYD 543 543
CC FT CARBOHYD 622 622
CC FT CARBOHYD 663 663
CC SQ SEQUENCE 877 AA; 96845 MW; 4418829ED871A249 CRC64;
Query Match 15.9%; Score 688.5; DB 1; Length 877;
Best Local Similarity 26.7%; Pred. No. 1,8e-34;
Matches 203; Conservative 126; Mismatches 343; Indels 87; Gaps 19;
Qy 126 FLQSKYEGSVRONSRGPKFLLVYNAITDLPATPNQGLYYQIVTQLP-----MINN 176
Db 7 FPEEDVYSAVLSRDVLEGGQLVNWKFSC-----NGKRAKYQESSRPADFVDEDMVYA 60
Qy 177 VMYFQINNKTKGALSTRBESQELNPAKNSYVLVSVKMGQSGSESPDITSVDIIV-- 234
Db 61 VNSFFPLSSHSKFLIYAQDKE-----TQEKQVAVKLSTKPALPEDSVKESREIEIIVP 115
Qy 235 -----TENIWKAPKPVENVNSTDPPIKITOVNRM-----DPCAQ 270
Db 116 ROYTKANGYLQROKRWVIR-PINLPENSRGPPQELVATIRSDRDNLSLRYSVTGPGA- 173
Qy 271 YSLVDREKLPPEPFSIDQ-EGDIYVTPQIDREKQAVYVAVAKDEYKPLSYPLEIHYK 329
Db 174 -----DPPPTGIFINPISGOLSTYKPLDRELLRPHLRRAHYDINGNOVENPIDIYIN 227
Qy 330 VKDINNPCTCSPVTVPEVOENERLGNISIGTLTAHDREENTANSFLNYRIVECTPKP 389
Db 228 VIDMNDNRPEFLHQWNSVPEGSRGTYVMVTVAIDADDPAALNGMLRYRLISQAFSTP 287
Qy 390 MDGLFLITVAGMLQAKOSLKKQOTPOYNLTIEVSDK-----FKLCEVQIVINDIN 443
Db 288 SPNPFITNNETGDIITVAAGLDREKVOQYTLTIQATDMEGNFTYGLSTATAVITVTVDN 347
Qy 444 DQIPIF-EKSDYGNLTAEADTNGSTLTITQATDDEPTGSSKILYHIKDGSEGLGVP 502
Db 348 DNPPEFTANTFYGEV---PENRVDIVANLTVDKQDPHTPMANAYRISGDPYGRFALIT 404
Qy 503 DTPDHTNGYVILIKKPLDFTAAVSNIVFKAENPEPLVFGVKYKNSAFKFTLITVDNE 562
Db 405 QTPDPSNDGLTVVAVRIPDETNRMFVLVAENQVPLAKGIQHPPOSTATVSVTVDVNE 464

```

QY 563 APOGSHVFOAKSSEDAVIGTKGNVTAKEDECI---DISYSLGDTGRLKIDHTVGT 619
 DB 465 NPFAPAPKRIROBEGHAGVLTFTFAODPDRYMOQINIRYTKLSDPANLTKIDSVNGOI 524
 QY 620 FSVAPIUDREA-----GSPYRVGVVATEVCGSSLSVSEPHILIMVNNPRLAKDYGLF 675
 DB 525 TTIAVLDRSEBPNVKNATYNATFELASDNGIIPMSGTGLQIYLLDINNAPOVPEQAEI 583
 QY 676 FCHPLSAPGSLIFEATDDOHLFRGPH-FTFSLGSGSLDMEVSKINGTHARLSTRHTD 734
 DB 584 -CE-TDPNLSINITALDYIDPDNAGPFAFDLPSPVTKKNTTRANGDPFAQLNKKIF 641
 QY 735 FEEBAYVVLIRINDGFRPPEGVSLPVTFCSC-VEGSCFRPAGHOTGIPTV---GNAV 790
 DB 642 LEAGIVEVPLIITDSGNPKSNISILRVKVCQCDNSGDC-----TDVDRIVAGLGTG 694
 QY 791 ILTLTLVIGIILAVFIRIKOKGKNVSAQASEK 829
 DB 695 AIIAIIILCIIITLILVIMFVMMKRRKEROAQLIDP 733

RESULT 13
 CAD2_RAT STANDARD; PRT; 906 AA.
 AC Q921Y3; Q9R0T5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).
 GN CMH2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley; TISSUE=Testis;
 RX MEDLINE=98187820; PubMed=9528971;
 RA Chung S.S., Mo M.Y., Silverstein B., Lee W.M., Cheng C.Y.;
 RT "Rat testicular N-cadherin: its complementary deoxyribonucleic acid
 cloning and regulation.";
 RL Endocrinology 139:1853-1862 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RA Asai K., Tada T., Yamamoto M., Tada A., Mizuno M., Himoto T., Kato T.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
 CC NEURONAL RECOGNITION MECHANISM.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
 CC -1- TISSUE SPECIFICITY: IN TESTIS, EXPRESSED IN SEROLI AND GERM
 CC CELLS.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF097593; AAC8318.1; -;
 CC EMBL; AB017695; BAA84919.1; -;
 CC HSSP; P5116; INCU.
 CC InterPro; IPR002126; Cadherin.
 CC InterPro; IPR000233; Cadherin_C-term.
 CC Pfam; PF00029; cadherin_5.
 CC Pfam; PF01049; cadherin_C-term; 1.
 CC PRINTS; PR00205; CADHERIN.

DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 3.
 DR PROSITE; PS0268; CADHERIN 2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KM Signal.
 FT 1 27
 FT 28 159
 FT 160 906
 FT 160 724
 FT 725 745
 FT 746 906
 FT 160 267
 FT 268 382
 FT 383 497
 FT 498 603
 FT 604 717
 FT 863 878
 FT 190 190
 FT 273 273
 FT 325 325
 FT 402 402
 FT 572 572
 FT 622 622
 FT 651 651
 FT 692 692
 FT 7 7
 FT 48 48
 FT 153 153
 FT 646 646
 FT 658 658
 FT 724 724
 SQ SEQUENCE 906 AA; 99685 MW; 97D9937ED8D1F8B5 CRC64;

Query Match 15.9%; Score 686.5; DB 1; Length 906;
 Best Local Similarity 27.2%; Pred. No. 2.5e-34;
 Matches 215; Conservative 126; Mismatches 344; Indels 105; Gaps 24;

QY 97 AALDANGIIV-----EGVPITIEVKDINDRPFGLSGKGVYKNSRPGKFLYVNA 150
 DB 21 ASLEHSGELALCKTGEPREYVSAVLPKYTHBQP-LLVNKFNSCNRK-----RKQYESS 74
 QY 151 TDLDDPATPNQGLYYQIVIQILPMINNVYFQI---NNKTG-----AISTRGSGOELNPA 202
 DB 75 EPADFKVDEDTIVY--AVASFPLSABQAKFLVLAQDKERQEKQVAVNLSLPSLUEEEM 132
 QY 203 KNP-SYNLIVSYKDM---GGQSENSFSDTTSVITITENIKAPKPEVEMENSTDPHPYK 258
 DB 133 KSPHEIEELVFPROLAKHSGALQROKRD-----WVLP-PINLPENSRGPFPOE 179
 QY 259 ITOVRNN-----DPGAQYSIVDKELPPFPFSIDQ-EGDIYVTOPLDREKKA 305
 DB 180 LVRIKSDRKSLSLKSVTGPQA-----DDPPGIFIPINISQSLSTKPLDRELIAR 232
 QY 306 YVFVAVADKYGKPLSYPLEIHKVYKVDINDNPCTPSPVTVVEVOENERLGSIGTLTAH 365
 DB 233 FHLRAVAVDINQVBNPIDIYINVIDNDNRPEFLHQVWNSVSEKSGRYWMTVIAI 292
 QY 366 DDEENTANSPLNTYIVBGTPEKLPMDGLFLITVAGMLQAKQSLKQDTPQYNTLIEVS 425
 DB 293 DADDPVALNGMRYRILSGAPSTPSNMFTINNEGDIITVAAGLDREKVOQYTLIQAT 352
 QY 426 DKD-----FKTLCEVOINVIDINDQIPF-EKSYGMLTLAEDNINISITLITQATAD 478
 DB 353 DNEGNPTYGSLSTATAVAVITVDVNNPPEFTMTTGYEV--PENRVAVIYANLVITTDQ 409
 QY 479 EPTFGSKLLVHIINGDESEGLVDTPHTNGYVLIKKPLDPETAASNVIFKAENEP 538
 DB 410 QHTPAAMNAAYRISGDPGRFALITDPNSNDGLVTVKPIDFTNRMFVLLVAENQVP 469
 QY 539 LVFGKYNAASSPAKTLTYTDVNEAPQSSQVHFAKXSEDAVIGTKVGNVTAKEDECI-- 596
 DB 470 IAKGIQHPPQSTATVATVIVDVENPYPAPNPKIIRQEBGHAGTMTLTTLTAODPDRYVQ 529

QY 597 -DISYSLRGDTRGWLKIDHTGELFVSAPLDRFA---GSPYRQVAVTEVSGSSLSVS 651
 DB 530 QNIRYTKLSDPANMLKIDPVNGQITTAVDRESPPNKNNIYATPLASDNGIPMSGTG 569
 QY 652 EPHILIMDVNDNPRILAKDYTGLEFFCHPLSA-----PQSLIFEATDDOHLFRGPH-F 703
 DB 590 TLGIYLLIDINDMNAQVL-----PQEAETCETPEPNSINTALDYIDIDNPAGPFAF 639
 QY 704 TFSLGSSSLQNDMEVSKINGTHARLSTRHDFEERAAVVLIRINDGRPPLEGVSLPYT 763
 DB 640 DLPLSPATIKRMNTITRLNDFQALNKIKFLEAGIYEVPIVITDSNPPKNSIILIRK 699
 QY 764 FCSC-VEGSCFRPAGHOTGIPTV---QMAVGILLTLTLVIGIILAVFIRIKDKGKNV 819
 DB 700 VCCDSNGDC-----TDVDRIVAGAGTGCTITAILICILILILVLMFVMMKRDK 752
 QY 820 ESNQASEVKP 829
 DB 753 RQAKQLIDP 762

RESULT 14
 CAD4 HUMAN STANDARD; PRT; 916 AA.
 ID CAD4 HUMAN STANDARD; PRT; 916 AA.
 AC P55283;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cadherin-4 precursor (Retinal-cadherin) (R-cadherin) (R-CAD).
 GN CDH4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95073006; PubMed=7982033;
 RA Tanihara H., Sano K., Heimerl R.U., St John T., Suzuki S.;
 RT "Cloning of five human cadherins clarifies characteristic features of
 RT structural extracellular domain and provides further evidence for two
 RT structurally different types of cadherin.";
 RL Cell Adhes. Commun. 2:15-26(1994).
 RN [2]
 RP SEQUENCE OF 393-916 FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=91283540; PubMed=2059658;
 RA Suzuki S., Sano K., Tanihara H.;
 RT "Diversity of the cadherin family: evidence for eight new cadherins
 RT in nervous tissue.";
 RL Cell Regul. 2:261-270(1991).
 RC Cell Regul. 2:261-270(1991).
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. MAY PLAY AN IMPORTANT ROLE
 CC IN RETINAL DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN BRAIN BUT ALSO FOUND IN
 CC OTHER TISSUES.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and that statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; L34059; AAA35627.1; -
 DR HSPD; P15116; INCI.
 DR Genew; HGNC:1763; CDH4.
 DR MIM; 603006; -.

DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR00233; Cadherin_C_term.
 DR Pfam; PF00028; cadherin; 5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 2.
 DR PROSITE; PS0268; CADHERIN 2; 5.
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
 KM Signal.
 FT SIGNAL. 1 20
 FT PROPEP 21 169
 FT CHAIN 170 916
 FT DOMAIN 170 734
 FT TRANSMEM 735 756
 FT DOMAIN 757 916
 FT DOMAIN 170 277
 FT DOMAIN 278 392
 FT DOMAIN 393 507
 FT DOMAIN 508 613
 FT DOMAIN 614 724
 FT DOMAIN 873 888
 FT CARBOHYD 283 283
 FT CARBOHYD 412 412
 FT CARBOHYD 557 557
 FT CARBOHYD 632 632
 FT CARBOHYD 661 661
 FT CARBOHYD 702 702
 SQ SEQUENCE 916 AA; 100446 MW; E717C54A19E0C52A CRC64;

Query Match 15.8%; Score 684.5; DB 1; Length 916;
 Best Local Similarity 29.3%; Pred. No. 3.4e-34;
 Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

QY 239 WKAPKEVENENSTDPHPIKITQVRW---NDPQAGVSL--VDKELPRPFPSTID-QEGDI 292
 DB 171 WVLP-PINVPENSRGFPPOOLVIRISDKONDPIRYSITGVGADQPMVEFISNSGMR 229
 QY 293 YVTQPLDREKKAIVYVAVAKDEYKPLSPLEIHVKVDINDNPTCPSPVTVFVQEN 352
 DB 230 YVTRPDREHASYHLRAHVDNNGKVENPIDLYIVDADNHEPFINQVNCVDEG 289
 QY 353 ERLGNSIGTLTAHDEENTANSFLNRYIVEQTPKLPMDGLFLIQYAGMLQAKOSLKK 412
 DB 290 SKRGTYVMTITANDADSTTANGMVRVRYLTQRPQSPQSMFTINSETGIVTVAGMDR 349
 QY 413 QDTPQYNLTIVESDKD-----FKTLCFQVQINVIDINDQIPPEKSDYGNLTAEPTNG 466
 DB 350 EKVQQTIVIVQATDMEGNLNGLSNTATATITVTDVNDPSEFTASTFAG--EVPENSV 407
 QY 467 STLTITQATDADDEPFGSSKILVHIKGSSEGLGVDTPHTNTGVVILKKPLDFETAAV 526
 DB 408 TVVANLTWMDRDPHPENNAVRIITSGDSGHFSVRIDPVTNEGAVTVKAVDYELNRA 467
 QY 527 SNIVFAENEPPELVFGKYVNASFAKFTLVTVDNENAPQFSQHVFAKVSDEVAIGTKVG 586
 DB 468 FMLTWVWSNQAPLASGIQMSFQSTAGVTISINDINABPFPSSHKLIRLEEGVPGLTVL 527
 QY 587 NVTAKDPEGL--DISYSLRGDTRGWLKIDHTGELFVSAPLDRFA---GSPYRQVAV 639
 DB 528 TFSALVDPRMQQAVYKSLSDPASWLNATNGQITTAVDRESIYTKNNVYEATFLA 587
 QY 640 TFSLGSSLSVSFPHILIMDVNDNPRILAKDYTGLEFFCHPLSAPSLIFEATDDOHLFR 699
 DB 588 ADNGIIPASGTGLQYLLIDINDNAPBLPKAQICERPNLN--INITADADVHPNI 644
 QY 700 GPH-FTFSLGSSSLQNDMEVSKINGTHARLSTRHDFEERAAVVLIRINDGRPPLEGV 758
 DB 645 GPYVFEIPFVPAVVRKMTITRLNGDYAQLSLAILVLEAGMVVPIIVTDSGNPPLSNTS 704
 QY 759 SLPTVFCSCVEGSCFRPAGHOTGIPTVGNAAVGLTLTLVIGIILAV--FIRIKDKGR 816
 DB 705 IIRKVCPCDNDGDCITIG---AVAAAGLGTGAIVAIIICILILITLVLLFVMMKRREK 761

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 20:21:49 ; Search time 8328 Seconds

(without alignments)
11689.347 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345

Sequence: 1 gaattccgctcgcacacatg.....cctttgcgcgcgaattc 3345

Scoring table: IDENTITY NUC
Gapop 10%0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_man:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3345	100.0	3345	6 AX193509	AX193509 Sequence
2	3345	100.0	3345	6 AX330272	AX330272 Sequence
3	3345	100.0	3345	6 AX332252	AX332252 Sequence
4	3345	100.0	3345	6 AX410642	AX410642 Sequence
5	3345	100.0	3345	6 HSU07969	U07969 Human Intes
6	3314.8	99.1	3654	6 AX330312	AX330312 Sequence
7	3314.8	99.1	3654	6 AX335495	AX335495 Sequence
8	3314.8	99.1	3654	6 AX411163	AX411163 Sequence
9	3314.8	99.1	3654	6 HSRNMLTFA	X03228 H.sapiens m
10	2499	74.7	2499	6 140307	140307 Sequence 2
11	2499	74.7	2499	6 181197	X78897 R.norvegicu
12	1774.4	53.0	3628	10 RNCADHER	D87912 Mus muscu
13	1758.6	52.6	3414	10 D87912	D87912 Mus muscu
14	1747.8	52.3	3351	10 AF177669	AF177669 Mus muscu
15	1184.8	35.4	1733	10 RA1EDRPT	L46874 Rattus norv
16	848.8	25.4	153894	9 AP003478	AP003478 Homo sapi
17	848.8	25.4	166450	9 AC021576	AC021576 Homo sapi
18	847.2	25.3	176854	9 AP003351	AP003351 Homo sapi
19	514.2	15.4	523	6 AX340772	AX340772 Sequence
20	508.8	15.2	1066	11 G26817	G26817 human STS S
21	497.2	14.9	501	6 AX341845	AX341845 Sequence
22	469.4	14.0	495	6 AX340324	AX340324 Sequence
23	461.6	13.8	492	6 AX339851	AX339851 Sequence
24	431.4	12.9	516	6 AX396194	AX396194 Sequence
25	410.4	12.3	568	6 AX396490	AX396490 Sequence
26	365.6	10.9	370	6 AX339931	AX339931 Sequence
27	342	10.2	342	6 AX3397080	AX3397080 Sequence
28	340.4	10.2	500	6 AX336419	AX336419 Sequence
29	331.8	9.9	335	6 AX260426	AX260426 Sequence
30	331.8	9.9	335	6 AX261891	AX261891 Sequence
31	331.4	9.9	346	6 AR166828	AR166828 Sequence
32	331.4	9.9	346	6 AX192454	AX192454 Sequence
33	319.8	9.6	325	6 AX193263	AX193263 Sequence
34	306	9.1	3373	5 AF428098	AF428098 Danio rer
35	269.6	8.1	274	6 AX339825	AX339825 Sequence
36	265.2	7.9	268	6 AX339794	AX339794 Sequence
37	265	7.9	265	6 AX351289	AX351289 Sequence
38	243.4	7.3	422	17 AG025648	AG025648 Homo sapi
39	211.6	6.3	119350	9 HS12803	Z98742 Human DNA S
40	210.6	6.3	85718	9 AC008438	AC008438 Homo sapi
41	210.6	6.3	147078	2 AC080184	AC080184 Homo sapi
42	210.6	6.3	174158	9 AC011399	AC011399 Homo sapi
43	210.6	6.3	259474	9 HUAC004605	AC004605 Homo sapi
44	210.2	6.3	229752	2 AC009994	AC009994 Homo sapi
45	209.6	6.3	129758	9 AC118269	AC118269 Homo sapi

ALIGNMENTS

RESULT 1

AX193509 3345 bp DNA linear PAT 15-AUG-2001

LOCUS AX193509

DEFINITION Sequence 1076 from Patent WO0149716.

ACCESSION AX193509

VERSION AX193509.1 GI:15211449

KEYWORDS

SOURCE

ORGANISM human.

Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y.

AUTHORS Compounds for immunotherapy and diagnosis of colon cancer and

TITLE

Pred. No. is the number of results predicted by chance to have a

Methods for their use
 Patent: WO 0149716-A 1076 12-JUL-2001;
 CORIXA CORPORATION (US);

FEATURES
 Source 1.3345
 Location/Qualifiers

BASE COUNT 951 a 746 c 720 g 928 t

ORIGIN

Query Match 100.0%; Score 3345; DB 6; Length 3345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAATTCCTGCTCGACCTGAAATGGAAGAAAGACTTTTAAACCACTTTTGTGACTTA 60
 Db 1 GAATTCCTGCTCGACCTGAAATGGAAGAAAGACTTTTAAACCACTTTTGTGACTTA 60
 Qy 61 CAGAAAGAAATTTGAATTAAGAAACTATGATACCTTCAGGCCATCTTCACTCCCTGTG 120
 Db 61 CAGAAAGAAATTTGAATTAAGAAACTATGATACCTTCAGGCCATCTTCACTCCCTGTG 120
 Qy 121 CTCTTATGCTTTATTTGGCACTGATATGCGCAAGGGGAAATTTAGTGAACCCCTG 180
 Db 121 CTCTTATGCTTTATTTGGCACTGATATGCGCAAGGGGAAATTTAGTGAACCCCTG 180
 Qy 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCACTTT 240
 Db 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCACTTT 240
 Qy 241 AAGCCATCTCTGCTGCTGACTTTTGAATTAAGTGGGAGACAGACAAATTTTGTG 300
 Db 241 AAGCCATCTCTGCTGCTGACTTTTGAATTAAGTGGGAGACAGACAAATTTTGTG 300
 Qy 301 ATGAAACGGAGGAGACTTCTGTATTAACAAGACCTTTGACAGGGGAAACAAGATCTACT 360
 Db 301 ATGAAACGGAGGAGACTTCTGTATTAACAAGACCTTTGACAGGGGAAACAAGATCTACT 360
 Qy 361 CACAAATCTCCAGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGAGGTCCTCCCT 420
 Db 361 CACAAATCTCCAGTTTGCAGCCCTGAGCGCTAATGGAATTAATGAGAGGTCCTCCCT 420
 Qy 421 ATCAACATGAAAGTGAAGACATCAACGACATGACCCAGTTTCTTCAATCAAGTAC 480
 Db 421 ATCAACATGAAAGTGAAGACATCAACGACATGACCCAGTTTCTTCAATCAAGTAC 480
 Qy 481 GAAGGCTCAGTAAGGACAGAACTTCGCCAGAGAAAGCCCTTCTGTATGTCATGCGACA 540
 Db 481 GAAGGCTCAGTAAGGACAGAACTTCGCCAGAGAAAGCCCTTCTGTATGTCATGCGACA 540
 Qy 541 GACCTGATGATTCGGGCACTCCCAATGCGAGCTTTATTAACAGATTTGATCAGCTT 600
 Db 541 GACCTGATGATTCGGGCACTCCCAATGCGAGCTTTATTAACAGATTTGATCAGCTT 600
 Qy 601 CCCATGATCAACATGTCATGATCTTTCAATCAACAAAGGGGAGCATCTCTT 660
 Db 601 CCCATGATCAACATGTCATGATCTTTCAATCAACAAAGGGGAGCATCTCTT 660
 Qy 661 ACCGAGAGGGATCTCAGAAATTAATCTCTTAAGAACTCTTCTTAATCTGCTGATC 720
 Db 661 ACCGAGAGGGATCTCAGAAATTAATCTCTTAAGAACTCTTCTTAATCTGCTGATC 720
 Qy 721 TCAATGAAAGACATGGAAGGCGAGAGTGAATTCCTTCAATGATCAACATCTGTGAT 780
 Db 721 TCAATGAAAGACATGGAAGGCGAGAGTGAATTCCTTCAATGATCAACATCTGTGAT 780
 Qy 781 ATCAATGACAGAGAAATTTGGAAGACCAAAACCTGTGAGATGTGAAACCTCA 840
 Db 781 ATCAATGACAGAGAAATTTGGAAGACCAAAACCTGTGAGATGTGAAACCTCA 840
 Qy 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGGTGCAATAT 900
 Db 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGGTGCAATAT 900

Qy 901 TCCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTTCCATTTTCAATGACAGAGAGAT 960
 Db 901 TCCTTAGTTGACAAAGAGAGAGCTGCCAAGATTCCTTCCATTTTCAATGACAGAGAGAT 960
 Qy 961 ATTTAGGTGACTCAGCCCTTGAACGAGAGAAAGAGATGATGTTTATTTATGCACTT 1020
 Db 961 ATTTAGGTGACTCAGCCCTTGAACGAGAGAAAGAGATGATGTTTATTTATGCACTT 1020
 Qy 1021 GCAAGAGATGATGAGGAAACCACTTCTATTCCTGCTGGAATTCATGTAAGTTAA 1080
 Db 1021 GCAAGAGATGATGAGGAAACCACTTCTATTCCTGCTGGAATTCATGTAAGTTAA 1080
 Qy 1081 GATATTATGATATATCACTTACATGTCGCTGACAGTAACTGATTTGAGTCCAGAG 1140
 Db 1081 GATATTATGATATATCACTTACATGTCGCTGACAGTAACTGATTTGAGTCCAGAG 1140
 Qy 1141 AATGAACGACTGGGTAAACAGTATCGGGAACCTTATGACATGACAGGGATGAAGAAAT 1200
 Db 1141 AATGAACGACTGGGTAAACAGTATCGGGAACCTTATGACATGACAGGGATGAAGAAAT 1200
 Qy 1201 ACTGCCAAGCTTTTCTAAACCTACAGATTTGAGAGCAACCTCCAACTTCCATGAT 1260
 Db 1201 ACTGCCAAGCTTTTCTAAACCTACAGATTTGAGAGCAACCTCCAACTTCCATGAT 1260
 Qy 1261 GGACTCTTCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTAAACAGTCTTGAAG 1320
 Db 1261 GGACTCTTCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTAAACAGTCTTGAAG 1320
 Qy 1321 AAGCAAGATATCTCTGACATCAACCTTAACAGATGAGGTGTCGACAAAGATTTCAAGACC 1380
 Db 1321 AAGCAAGATATCTCTGACATCAACCTTAACAGATGAGGTGTCGACAAAGATTTCAAGACC 1380
 Qy 1381 CTGTGTTTGTGCAATCAACCTTATGATTAATCAAGATTCCTTCAATTTGAAATA 1440
 Db 1381 CTGTGTTTGTGCAATCAACCTTATGATTAATCAAGATTCCTTCAATTTGAAATA 1440
 Qy 1441 TCAGATTTAGAAACCTGACTCTTGTGTAAGACACAAACCTTGGGTCCACATCTTAAC 1500
 Db 1441 TCAGATTTAGAAACCTGACTCTTGTGTAAGACACAAACCTTGGGTCCACATCTTAAC 1500
 Qy 1501 ATCCAGGCACTGATGCTGATGAGCATTTTACGAGGATCTTAATTTCTGATCATATTC 1560
 Db 1501 ATCCAGGCACTGATGCTGATGAGCATTTTACGAGGATCTTAATTTCTGATCATATTC 1560
 Qy 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATACCAACCGGA 1620
 Db 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATACCAACCGGA 1620
 Qy 1621 TATGTCATTAATAAAGCCTTGTGATTTGAAACAGCAGCTGTTTCCACATTTGTTTC 1680
 Db 1621 TATGTCATTAATAAAGCCTTGTGATTTGAAACAGCAGCTGTTTCCACATTTGTTTC 1680
 Qy 1681 AAAGCAGAAATCTCTAGCCTTGTGATTTGATGTAAGTACAAATGCAAGTCTTTTGGC 1740
 Db 1681 AAAGCAGAAATCTCTAGCCTTGTGATTTGATGTAAGTACAAATGCAAGTCTTTTGGC 1740
 Qy 1741 AAGTTCAAGCTTATTTGACAGATGTAAGAGACCTCAATTTTCCAAACAGTATTC 1800
 Db 1741 AAGTTCAAGCTTATTTGACAGATGTAAGAGACCTCAATTTTCCAAACAGTATTC 1800
 Qy 1801 CAAGCAGAAATCTGAGAGATGATGCTATGAGCACTTAATGAGGCAATGTGACCTCAAG 1860
 Db 1801 CAAGCAGAAATCTGAGAGATGATGCTATGAGCACTTAATGAGGCAATGTGACCTCAAG 1860
 Qy 1861 GATCCAGAAAGTCTGACATTAAGCTATTCATGAGGGAGACACAAAGGTGGCTTTAA 1920
 Db 1861 GATCCAGAAAGTCTGACATTAAGCTATTCATGAGGGAGACACAAAGGTGGCTTTAA 1920
 Qy 1921 ATTGACCAAGTGTGAGTGTGATCTTTAATGTGCTTCATTTGACAGAGAAAGCGGAAGT 1980
 Db 1921 ATTGACCAAGTGTGAGTGTGATCTTTAATGTGCTTCATTTGACAGAGAAAGCGGAAGT 1980

QY	1981	CCATATCCGGGTAACAAGTGTGGCCACAGAAAGAGGGAGGTCTTCCTTAAGCTCTGTGCA	2048
Db	1981	CCATATCCGGGTAACAAGTGTGGCCACAGAAAGAGGGAGGTCTTCCTTAAGCTCTGTGCA	2048
QY	2041	GAGTTCACCTGCATCTTAATGAGATGATGACAAACCTCCAGGCTAGCCAAAGACTAC	2100
Db	2041	GAGTTCACCTGCATCTTAATGAGATGATGACAAACCTCCAGGCTAGCCAAAGACTAC	2100
QY	2101	AAGGCTTGTTCTTCTGTGCATCCCTCAGTGCACCTGGAAAGTCTCATTTTGGAGCTACT	2160
Db	2101	AAGGCTTGTTCTTCTGTGCATCCCTCAGTGCACCTGGAAAGTCTCATTTTGGAGCTACT	2160
QY	2161	GATGATGATCAGCACTTAATTCGGGGGTCCCAATTTAATTTCCCTGGCAGTGGAAAGC	2220
Db	2161	GATGATGATCAGCACTTAATTCGGGGGTCCCAATTTAATTTCCCTGGCAGTGGAAAGC	2220
QY	2221	TTACAAAAGACTGGGAAGTTTCCAAAATCAATGTACTCATGCCCCA.CTGTCTTACAGAG	2280
Db	2221	TTACAAAAGACTGGGAAGTTTCCAAAATCAATGTA.CTCAATGCCAGCTGTCTTACAGAG	2280
QY	2281	CACACAGACTTTGAGGAAAGGGCGTATGTCGTCTTGATCCGATCATGATGGGGGTCCG	2340
Db	2281	CACACAGACTTTGAGGAAAGGGCGATGTCGTCTTGATCCCAATCATGATGGGGGTCCG	2340
QY	2341	CAACCCCTGGAAAGGATGTGTTCTTTTACCAATTCATCTCAGTTCAGTGTGTGTGGAAAGT	2400
Db	2341	CAACCCCTGGAAAGGATGTGTTCTTTTACCAATTCATCTCAGTGTGTGTGGAAAGAGT	2400
QY	2401	TGTTTCCGGCAGCAGAGTCAACACAGCTGGATACCACTGTGGGCAATGGCAGTTGGTATA	2460
Db	2401	TGTTTCCGGCAGCAGAGTCAACAGCTGGATACCACTGTGGGCAATGGCAGTTGGTATA	2460
QY	2461	CTGCTGACCAACCCCTTCTGTGATTTGTATTAATTTTACAGTGTGTGTTATCCGAAATAAG	2520
Db	2461	CTGCTGACCAACCCCTTCTGTGATTTGTATTAATTTTACAGTGTGTGTTATCCGAAATAAG	2520
QY	2521	AAGGTAAGGCAAAAGATAATGTTGAAAGTCTCAAGACTTGAAGTCAAACTCTGAGA	2580
Db	2521	AAGGTAATAGGCAAAAGATAATGTTGAAAGTCTCAAGACTTGAAGTCAAACTCTGAGA	2580
QY	2581	AGCTGAATTTAAAAAGAAATGTTGAATTTATATAGCAAGTGCATTTACAGACAACA	2640
Db	2581	AGCTGAATTTAAAAAGAAATGTTGAATTTATATATAGCAAGTGCATTTACAGACAACA	2640
QY	2641	TCTCATCTTAATCTTTCAATCTAACGTGACATTAATTTTAAACAGATATTCCTCT	2700
Db	2641	TCTCATCTTAATCTTTCAATCTAACGTGACATTAATTTTAAACAGATATTCCTCT	2700
QY	2701	TGTCCTTTAATTTTGTCTTAATTTCTTTTGTAGGTGAGTCTTGCTGTGTGCCAG	2760
Db	2701	TGTCCTTTAATTTTGTCTTAATTTCTTTTGTAGGTGAGTCTTGCTGTGTGCCAG	2760
QY	2761	GCTGAGATGACAGTGTGTGATCCAGCTCACTGAAACCTCCGCGCTCCAGGGTTCAATGA	2820
Db	2761	GCTGAGATGACAGTGTGTGATCCAGCTCACTGAAACCTCCGCGCTCCAGGGTTCAATGA	2820
QY	2821	TTCTCTGCTCAGCTTCTTAAGTACGTGGGTTTACAGGCAACCAACACATGCCCCAGCT	2880
Db	2821	TTCTCTGCTCAGCTTCTTAAGTACGTGGGTTTACAGGCAACCAACATGCCCCAGCT	2880
QY	2881	AATTTTGTATTTTAAATATAGAGCGGGGTTTCCCAATTTGGCCAGGCTGTCTTGAATCTC	2940
Db	2881	AATTTTGTATTTTAAATATAGAGCGGGGTTTCCCAATTTGGCCAGGCTGTCTTGAATCTC	2940
QY	2941	CTGACGTCAGATGATCTGCCCTGCTGCCATTAACAGGCAATGAACCACTGCAACCA	3000
Db	2941	CTGACGTCAGATGATCTGCCCTGCTGCCATTAACAGGCAATGAACCACTGCAACCA	3000
QY	3001	CTTACTTAAGATATTTCAATGTGCTATAGACATTAAGAGATTTTTCATTTTCCATGACAT	3060
Db	3001	CTTACTTAAGATATTTCAATGTGCTATATAGACATTAAGAGATTTTTCATTTTCCATGACAT	3060
QY	3061	TTTTTCTCTCTGCAAAATGAGCTTAAGTACTTGTGTTTTCCCTTTTGGGCAAGACAGACT	3120

Db	3061	TTTTCTCTCGGAAATGGCTTACTACTGTGTTTTCCCTTTTGGGGCAAGCAGACT	3120
OY	3121	CATTAAATATCTCTACATTTTTCTTATTCAGAAGATATATCAGTGTGTCTCATAGA	3180
Db	3121	CATTAAATATCTGTACATTTTCTTATTCAGAAGATATATCAGTGTGTCTCATAGA	3180
OY	3181	ACTGCTTGAGATTGCATTTATGTTTTTCTGTGATTCATCCTGTGTGCCCTTCATCCTTGAC	3240
Db	3181	ACTGCTTGAGATTGCATTTATGTTTTTCTGTGATTCATCCTGTGTGCCCTTCATCCTTGAC	3240
OY	3241	TCCTTTGGTATTTCACTGAATTTCAACATTTGCAGAGAAAGAAAAAGTAGAGACTCAG	3300
Db	3241	TCCTTTGGTATTTCACTGAATTTCAACATTTGTGCAGAGAAAGAAAAAGTAGAGACTCAG	3300
OY	3301	GAAAAATTAATTAATTAAGAAAGACGCTTTTGGGCGCGGCAATTC	3345
Db	3301	GAAAAATTAATTAATTAAGAAAGACGCTTTTGGGCGCGGCAATTC	3345

RESULT 2	AX330272	3345 bp	DNA	linear	PAT 09-JAN-2002
LOCUS	AX330272	781 from Patent WO0194629.			
DEFINITION	Sequence				
ACCESSION	AX330272				
VERSION	AX330272.1	GI:18103250			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G., Horrigan, S., Soppet, D. R. and Weaver, Z.				
TITLE	Cancer gene determination and therapeutic screening using signature gene sets				
JOURNAL	Patent: WO 0194629-A 781 13-DEC-2001; Avalon Pharmaceuticals (US)				
FEATURES	Location/Qualifiers				
source	1..3345				
BASE COUNT	/organism="Homo sapiens" /db_xref="taxon:9606"				
ORIGIN	951 a 746 c 720 g 928 t				
Query Match	100.0%; Score 3345; DB 6; Length 3345;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 3345; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
QY	1 GAATTCGCTCCGACCACTGAATGGAAGAAAGACCTTTTAACCAACATTTGTGACTTA	60			
Db	1 GAATTCGCTCCGACCACTGAATGGAAGAAAGACCTTTTAACCAACATTTGTGACTTA	60			
QY	61 CAGAAAGAAATTGAATAAGAAAACATATGATCTTCAGGCCATCTTCACTCCCTGTGT	120			
Db	61 CAGAAAGAAATTGAATAAGAAAACATATGATCTTCAAGCCATCTTCACTCCCTGTGT	120			
QY	121 CTCTTATGCTTTAATTTGGCACTGATATGGCCCAAGAGGGAGATTAGTGACCCCTG	180			
Db	121 CTCTTATGCTTTAATTTGGCACTGATATGGCCCAAGAGGGAGATTAGTGACCCCTG	180			
QY	181 AAAACCCATGACATTTTCTATTATTAAGAAGCCCAAGAACCGAGTCMAATATATATTCAGATT	240			
Db	181 AAAACCCATGACATTTTCTATTATTAAGAAGCCCAAGAACCGAGTCMAATATATTCAGATT	240			
QY	241 AAGGCCAATCTCTCTGCTGTGACTTTTGAACATACTGGGAGACAGACAATATTTGTG	300			
Db	241 AAGGCCAATCTCTCTGCTGTGACTTTTGAACATACTGGGAGAGACAGACAATATTTGTG	300			
QY	301 ATGGAACGGGAGGAGACTTCGTATTTCAACAGAGCCCTGGACAGGGAAACAAGATCTACT	360			
Db	301 ATGGAACGGGAGGAGACTTCGTATTTCAACAGAGCCCTGGACAGGGAAACAAGATCTACT	360			

OY	361	CACAAATCTCCAGGTTGCGAGCCCTGAGCCCTAAATGGAAATTTATAGGAGGGGTCCAGTCCCT	420
Db	361	CACAAATCTCCAGGTTGCGAGCCCTGAGCCCTAAATGGAAATTTATAGGAGGGGTCCAGTCCCT	420
OY	421	ATCAACCAATAGAAGTGAAGGAGCATCAACGACAAATGAGCCACGTTCTTCCAGTCAAGATAC	480
Db	421	ATCAACCAATAGAAGTGAAGGAGCATCAACGACAAATGAGCCACGTTCTTCCAGTCAAGATAC	480
OY	481	GAAAGCTCAGTAAGGCAGAACTCTCGCCGAGAAAGCCCTTCTTGATATGTCATGTCACA	540
Db	481	GAAAGCTCAGTAAGGCAGAACTCTCGCCGAGAAAGCCCTTCTTGATATGTCATGTCACA	540
OY	541	GACCTGGATGATTCGGGCCATCTCCCAATGGCCAGCTTTATACGAGATTGTCAATCCACACTT	600
Db	541	GACCTGGATGATTCGGGCCATCTCCCAATGGCCAGCTTTATACGAGATTGTCAATCCACACTT	600
OY	601	CCCAATGATCAACAATGATCATGTACTTGTAGATCAACAACAAAGGAGGACATCTCTCTT	660
Db	601	CCCAATGATCAACAATGATCATGTACTTGTAGATCAACAACAAAGGAGGACATCTCTCTT	660
OY	661	ACCCGAGAGGAGTCTCAGAAATGGATCTGTAGAAATCCTTCTTAATCTGTGTATC	720
Db	661	ACCCGAGAGGAGTCTCAGAAATGGATCTGTAGAAATCCTTCTTAATCTGTGTATC	720
OY	721	TCAGTGAAGGACATGGGAGGCCAGAGTGAAGATTCCTTCACTGATATACCAATCTGTGTATC	780
Db	721	TCAGTGAAGGACATGGGAGGCCAGAGTGAAGATTCCTTCACTGATATACCAATCTGTGTATC	780
OY	781	ATCATATGACACAGAGAAATTTTGGAAACACCAAAACCTGTGAGATGTTGGAAATCTCA	840
Db	781	ATCATATGACACAGAGAAATTTTGGAAACACCAAAACCTGTGAGATGTTGGAAATCTCA	840
OY	841	ACTGATCTCTCAACCCCATCAAAATCACTCAGGTGCGGTGGAAATGATCCCGGTGCAATAT	900
Db	841	ACTGATCTCTCAACCCCATCAAAATCACTCAGGTGCGGTGGAAATGATCCCGGTGCAATAT	900
OY	901	TCCTTATGTTGACAAAGAGAGAGTGTGCAAGATTTCCATTTTCAATTTGACAGGAAAGGAT	960
Db	901	TCCTTATGTTGACAAAGAGAGAGTGTGCAAGATTTCCATTTTCAATTTGACAGGAAAGGAT	960
OY	961	ATTTACGTGACTCAGCCCTTGGACCGAGAAAGAAAGATGCAATATGTTTTTATGACAGTT	1020
Db	961	ATTTACGTGACTCAGCCCTTGGACCGAGAAAGAAAGATGCAATATGTTTTTATGACAGTT	1020
OY	1021	GCAAGAGTGAAGTACGAAAAACAATTTCATATCCGCTGAGAAATTCATGTAAAAAGTTAAA	1080
Db	1021	GCAAGAGTGAAGTACGAAAAACAATTTCATATCCGCTGAGAAATTCATGTAAAAAGTTAAA	1080
OY	1081	GATATTAATGATATATCCATCATATGTCCTGACCAAGTAAACCGATTTGAAGTCCAGAG	1140
Db	1081	GATATTAATGATATATCCATCATATGTCCTGACCAAGTAAACCGATTTGAAGTCCAGAG	1140
OY	1141	AATGAACGATCGGGGTAACAATATCGGGACCTTATGCAATGACAGGGATGAAAGAAAT	1200
Db	1141	AATGAACGATCGGGGTAACAATATCGGGACCTTATGCAATGACAGGGATGAAAGAAAT	1200
OY	1201	ACTGCAACAGTTTTCTAAATCTACAGATTTGTGAGCAAACTCCAAACTTCCATGGAT	1260
Db	1201	ACTGCAACAGTTTTCTAAATCTACAGATTTGTGAGCAAACTCCAAACTTCCATGGAT	1260
OY	1261	GGACTCTTCTAATCCAAACCTATGCTGGAATGTACAGTTAGCTTAAACAGTCTCTGAG	1320
Db	1261	GGACTCTTCTAATCCAAACCTATGCTGGAATGTACAGTTAGCTTAAACAGTCTCTGAG	1320
OY	1321	AAGCAAGATACCTTCAGTACAACTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACC	1380
Db	1321	AAGCAAGATACCTTCAGTACAACTTAACGATAGAGGTGTCTGACAAAGATTTCAAGACC	1380
OY	1381	CTTTGTTTGTCAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA	1440
Db	1381	CTTTGTTTGTCAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA	1440
OY	1441	TCAGATTAATGAAACCTGACTCTTGCTGAGACACAAACATTTGGTCCACATCTTAAAC	1500

[illegible]

Db 2521 AAGGATAAGGCAAGATTAATGTTGAAAGTGCTCAAGCATCTGAATGCAAACTCTGAGA 2580
Qy 2581 AGCTGAATTTGAAAGAGATGTTGAATTTATATAGCAAGTCTATTTCAGCAACA 2640
Db 2581 AGCTGAATTTGAAAGAGATGTTGAATTTATATAGCAAGTCTATTTCAGCAACA 2640
Qy 2641 TCTCATCTTATTTACTTTTCACTTACGTCATTAATTTTAAACAGATATTCCTCT 2700
Db 2641 TCTCATCTTATTTACTTTTCACTTACGTCATTAATTTTAAACAGATATTCCTCT 2700
Qy 2701 TGTCTTAAATTTATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTCTCTGCTCCAG 2760
Db 2701 TGTCTTAAATTTATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTCTCTGCTCCAG 2760
Qy 2761 GCTGAGATGAGTGTGTGATCCAGCTCACTGCAACCTCGCTCTGAGGTTTCACTGA 2820
Db 2761 GCTGAGATGAGTGTGTGATCCAGCTCACTGCAACCTCGCTCTGAGGTTTCACTGA 2820
Qy 2821 TTTCTCTGCTCAAGCTTCTTAAGTGTGAGTGTGAGGCAACCAACCAATGCTCCAGCT 2880
Db 2821 TTTCTCTGCTCAAGCTTCTTAAGTGTGAGTGTGAGGCAACCAACCAATGCTCCAGCT 2880
Qy 2881 AATTTTGTATTTTATATAGAGAGGGGTTTCGCAATTTGGCCAGGCTGCTTTGAATCT 2940
Db 2881 AATTTTGTATTTTATATAGAGAGGGGTTTCGCAATTTGGCCAGGCTGCTTTGAATCT 2940
Qy 2941 CTGACGTCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Db 2941 CTGACGTCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3000
Qy 3001 CTAATCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3060
Db 3001 CTAATCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3060
Qy 3061 TTTTCTCTCTGCAAAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Db 3061 TTTTCTCTCTGCAAAATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3120
Qy 3121 CATTAATATTTCTGATCATTTTCTTATCAAGAGATATATCAAGTGTGCTCATAGA 3180
Db 3121 CATTAATATTTCTGATCATTTTCTTATCAAGAGATATATCAAGTGTGCTCATAGA 3180
Qy 3181 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3240
Db 3181 ACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3240
Qy 3241 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
Db 3241 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3300
Qy 3301 GAAAAATTAATTAATAAAGACAGCCTTTGGGGCCGGAATTC 3345
Db 3301 GAAAAATTAATTAATAAAGACAGCCTTTGGGGCCGGAATTC 3345

RESULT 3
AX332252 3345 bp DNA linear PAT 09-JAN-2002
LOCUS AX332252
DEFINITION Sequence 2761 from Patent WO0194629.
ACCESSION AX332252
VERSION AX332252.1 GI:18122886
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endres, G.,
Horizgan, S., Soppet, D. R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 2761 13-DEC-2001;
Avallon Pharmaceuticals (US)

FEATURES
source 1.3345
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 951 a 746 c 720 g 928 t
ORIGIN
Query Match 100.0%; Score 3345; DB 6; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAAATCGCTCGACCACTGAATGAAAGAAAGAACTTTTAAACCAATTTTGTGACTTA 60
Db 1 GAAATCGCTCGACCACTGAATGAAAGAAAGAACTTTTAAACCAATTTTGTGACTTA 60
Qy 61 CAGAAAGAAATTTGAAATTAAGAAACTATGATCTTCAAGCCCATCTTCACTCCCTGTGT 120
Db 61 CAGAAAGAAATTTGAAATTAAGAAACTATGATCTTCAAGCCCATCTTCACTCCCTGTGT 120
Qy 121 CTTCTATGCTTATTTGGCACTGATATGCGCAAGAGGGGAAATTTAGTGAACCCCTG 180
Db 121 CTTCTATGCTTATTTGGCACTGATATGCGCAAGAGGGGAAATTTAGTGAACCCCTG 180
Qy 181 AAACCCATGACATTTTCTATTTATGAAGCCAAAGAACCAATTTATTTCAATTT 240
Db 181 AAACCCATGACATTTTCTATTTATGAAGCCAAAGAACCAATTTATTTCAATTT 240
Qy 241 AAGGCCAATCTCTCTGCTGATCTTGAATCTATGCGGGAGACAGACAAATTTTGTG 300
Db 241 AAGGCCAATCTCTCTGCTGATCTTGAATCTATGCGGGAGACAGACAAATTTTGTG 300
Qy 301 ATGAAAGGAGAGGACCTTCTGATTAACAAGAGCTTGGACAGGAAACAAGATCTACT 360
Db 301 ATGAAAGGAGAGGACCTTCTGATTAACAAGAGCTTGGACAGGAAACAAGATCTACT 360
Qy 361 CAATATCTCAGGTTGACAGCCCTGACGCTTAATGAATTAATGAGGGTCCAGTCCCT 420
Db 361 CAATATCTCAGGTTGACAGCCCTGACGCTTAATGAATTAATGAGGGTCCAGTCCCT 420
Qy 421 ATACCATGATGATGAGGACATCAACGACATGACCCAGTTTCTCCAGTAAAGTAC 480
Db 421 ATACCATGATGATGAGGACATCAACGACATGACCCAGTTTCTCCAGTAAAGTAC 480
Qy 481 GAAGGCTCAGTAAAGGAGAACTGCGCGAGAGAAAGCCCTTCTGATGATGATGACACA 540
Db 481 GAAGGCTCAGTAAAGGAGAACTGCGCGAGAGAAAGCCCTTCTGATGATGATGACACA 540
Qy 541 GACCTGATGATCCGCGCACTCCCAATGCGCAAGCTTTATTAACAGATTTGATCAGCTT 600
Db 541 GACCTGATGATCCGCGCACTCCCAATGCGCAAGCTTTATTAACAGATTTGATCAGCTT 600
Qy 601 CCCATATCAACATCTCATGTAATTTTCAAGATCAACAACAAACGAGGCCATCTCTTT 660
Db 601 CCCATATCAACATCTCATGTAATTTTCAAGATCAACAACAAACGAGGCCATCTCTTT 660
Qy 661 ACCCGAGAGGATCTCAGGAATTTGAATCTGCTAAGAAATCTTCCATATCTGAGATC 720
Db 661 ACCCGAGAGGATCTCAGGAATTTGAATCTGCTAAGAAATCTTCCATATCTGAGATC 720
Qy 721 TCAAGTAAAGACATGAGGAGCCAGAGTGAATTCCTTCAAGTATCAACATCTGAGAT 780
Db 721 TCAAGTAAAGACATGAGGAGCCAGAGTGAATTCCTTCAAGTATCAACATCTGAGAT 780
Qy 781 ATCATATGATACAGAGATTTTGAAGAGCAACAAACCTGTGAGATGAGAAACCTCA 840
Db 781 ATCATATGATACAGAGATTTTGAAGAGCAACAAACCTGTGAGATGAGAAACCTCA 840
Qy 841 ACTGATCTCAGCCCATCAAAATCACTCAGTGGGTGAGATGATCCGGTGCACAAATAT 900
Db 841 ACTGATCTCAGCCCATCAAAATCACTCAGTGGGTGAGATGATCCGGTGCACAAATAT 900
Qy 901 TCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTATTTCAATTGACAGAGAGAGAT 960
Db 901 TCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTATTTCAATTGACAGAGAGAGAT 960

Db 901 TCCTTAGTGA CAAGAAGAGCTGCCAAGATTCOCATTTTCAATGACAGAGAGAT 960
Qy 961 ATTTACGTGAC TACGCCCTTGGACCGAGAGAAAGATGCAATGTTTTTATGCAATT 1020
Db 961 ATTTACGTGAC TACGCCCTTGGACCGAGAGAAAGATGCAATGTTTTTATGCAATT 1020
Qy 1021 GCAAGAGATGAGTACGAGAAACACCTTCATATCCGTGAAATTCATGTAAGTAA 1080
Db 1021 GCAAGAGATGAGTACGAGAAACACCTTCATATCCGTGAAATTCATGTAAGTAA 1080
Qy 1081 GATATTAATGATATCCACCTACATGTCGTCACAGTAAACGTAATTTGAAGTCCAGAG 1140
Db 1081 GATATTAATGATATCCACCTACATGTCGTCACAGTAAACGTAATTTGAAGTCCAGAG 1140
Qy 1141 AATGACGACTGGGTAAACAGTATGGGACCTTATCTGCAATGACAGGATGAAAGAAAT 1200
Db 1141 AATGACGACTGGGTAAACAGTATGGGACCTTATCTGCAATGACAGGATGAAAGAAAT 1200
Qy 1201 ACTGCCAAGTCTTCTTAACCTACAGATTTGGAGCAAACTCCCAACTTCCATGGAT 1260
Db 1201 ACTGCCAAGTCTTCTTAACCTACAGATTTGGAGCAAACTCCCAACTTCCATGGAT 1260
Qy 1261 GGACTCTTCTTAATCCAAACCTATGCTGAAATGTTACAGTTAGCTAAACAGTCTTGAAG 1320
Db 1261 GGACTCTTCTTAATCCAAACCTATGCTGAAATGTTACAGTTAGCTAAACAGTCTTGAAG 1320
Qy 1321 AAGCAAGATCTCTCTGATACAACTTAAACGATAGAGTGTCTGACAAAGATTTCAAGACC 1380
Db 1321 AAGCAAGATCTCTCTGATACAACTTAAACGATAGAGTGTCTGACAAAGATTTCAAGACC 1380
Qy 1381 CTTTGTGTTGGCAAACTCAACGTTATGATCAATGATGAGATCCCATCTTTGAAAA 1440
Db 1381 CTTTGTGTTGGCAAACTCAACGTTATGATCAATGATGAGATCCCATCTTTGAAAA 1440
Qy 1441 TCAGATTATGAGAACTGACTCTTCTGAGAGACAAACTGGGTCACACTTTAAC 1500
Db 1441 TCAGATTATGAGAACTGACTCTTCTGAGAGACAAACTGGGTCACACTTTAAC 1500
Qy 1501 ATCCAGGCCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATATC 1560
Db 1501 ATCCAGGCCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATATC 1560
Qy 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATTAACACCGGA 1620
Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATTAACACCGGA 1620
Qy 1621 TATGTCAATTAATAAAAGCCTCTTGAATTTGAAAACAGAGCTGTTTCCAACTTGTCTC 1680
Db 1621 TATGTCAATTAATAAAAGCCTCTTGAATTTGAAAACAGAGCTGTTTCCAACTTGTCTC 1680
Qy 1681 AAAGCAGAAAATCTGAGCCTCTGATGTTTGTGTGAAGTACATGCAAGTCTTTGGCC 1740
Db 1681 AAAGCAGAAAATCTGAGCCTCTGATGTTTGTGTGAAGTACATGCAAGTCTTTGGCC 1740
Qy 1741 AAGTTCAAGCTTATTTGAGACAGATGTAAGAGACCTCAATTTTCCCAACAGTATTC 1800
Db 1741 AAGTTCAAGCTTATTTGAGACAGATGTAAGAGACCTCAATTTTCCCAACAGTATTC 1800
Qy 1801 CAAGCGAAAGTCACTGAGAGATGTAAGGCACTAAAGTGGGCAATGTGACGCCAAG 1860
Db 1801 CAAGCGAAAGTCACTGAGAGATGTAAGGCACTAAAGTGGGCAATGTGACGCCAAG 1860
Qy 1861 GATCCAGAGGTCGACATAGCTATTCATGAGGGGAGACAAAGAGTGTGCTTAA 1920
Db 1861 GATCCAGAGGTCGACATAGCTATTCATGAGGGGAGACAAAGAGTGTGCTTAA 1920
Qy 1921 AATGACACGCTGATGAGATCTTTAGTGTGCTTCAATTTGACAGAGAGCGGAAGT 1980
Db 1921 AATGACACGCTGATGAGATCTTTAGTGTGCTTCAATTTGACAGAGAGCGGAAGT 1980
Qy 1981 CCATATCGGATACAAAGTGTGGCCACAGAAATAGGGGGGCTTCTTAAAGCTGTGTCA 2040
Db 1981 CCATATCGGATACAAAGTGTGGCCACAGAAATAGGGGGGCTTCTTAAAGCTGTGTCA 2040

Qy 2041 GAGTTCACCTGATCCTTATGATGATGATGACAAACCTCCAGGCTAGCAAGACTAC 2100
Db 2041 GAGTTCACCTGATCCTTATGATGATGATGATGACAAACCTCCAGGCTAGCAAGACTAC 2100
Qy 2101 ACGGGCTTGTCTTCTGACCATCCCTCACTGACCTGAAAGTCTCATTTTGAAGCTACT 2160
Db 2101 ACGGGCTTGTCTTCTGACCATCCCTCACTGACCTGAAAGTCTCATTTTGAAGCTACT 2160
Qy 2161 GATGATGATCAGACTTATTTGGGGTCCCATTTTACATTTTCCCTGGAGTGAAG 2220
Db 2161 GATGATGATCAGACTTATTTGGGGTCCCATTTTACATTTTCCCTGGAGTGAAG 2220
Qy 2221 TTAACAAAGACGAGGAGGTTTCCAAATCAATGTAATGTAATGTAATGTAATGTAATG 2280
Db 2221 TTAACAAAGACGAGGAGGTTTCCAAATCAATGTAATGTAATGTAATGTAATGTAATG 2280
Qy 2281 CACACAGACTTGAAGAGAGGGGATGTCGTTTATGATCCGATCAATGATGAGGGGTCCG 2340
Db 2281 CACACAGACTTGAAGAGAGGGGATGTCGTTTATGATCCGATCAATGATGAGGGGTCCG 2340
Qy 2341 CCACCTTGGAGAGCATGTTTCTTTTACAGTTACATTCGACGTTGAGAGAGAGT 2400
Db 2341 CCACCTTGGAGAGCATGTTTCTTTTACAGTTACATTCGACGTTGAGAGAGAGT 2400
Qy 2401 TGTTCGCGCAGCAGGATCAGCAGACTGGGATTAACCACTGGGGCAGTGGATTA 2460
Db 2401 TGTTCGCGCAGCAGGATCAGCAGACTGGGATTAACCACTGGGGCAGTGGATTA 2460
Qy 2461 CTGCTGACCAACCTTCTGATGATGATTAATTTAGAGTGTGTTATCCCATTAAG 2520
Db 2461 CTGCTGACCAACCTTCTGATGATGATTAATTTAGAGTGTGTTATCCCATTAAG 2520
Qy 2521 AAGATTAAGGCAAAATATGTTAAAGTCTCAAGCATCTGAAGTCAAACTCTGAG 2580
Db 2521 AAGATTAAGGCAAAATATGTTAAAGTCTCAAGCATCTGAAGTCAAACTCTGAG 2580
Qy 2581 AAGATTAAGGCAAAATATGTTAAAGTCTCAAGCATCTGAAGTCAAACTCTGAG 2640
Db 2581 AAGATTAAGGCAAAATATGTTAAAGTCTCAAGCATCTGAAGTCAAACTCTGAG 2640
Qy 2641 TCTCATCTTAATTAATTTTCACTAAGTGTGATTAATTTTAAACAGATATTCCTCT 2700
Db 2641 TCTCATCTTAATTAATTTTCACTAAGTGTGATTAATTTTAAACAGATATTCCTCT 2700
Qy 2701 TGTCTTAAATTTTGTAAATTTCTTTTGAAGTGAAGTCTGTCGTGCGCCAG 2760
Db 2701 TGTCTTAAATTTTGTAAATTTCTTTTGAAGTGAAGTCTGTCGTGCGCCAG 2760
Qy 2761 GGTGAGTACAGTGTGATGTCAGCTCACTCACTCCGCTCTGGGTTCAATGA 2820
Db 2761 GGTGAGTACAGTGTGATGTCAGCTCACTCACTCCGCTCTGGGTTCAATGA 2820
Qy 2821 TTTCTGCTCAGCTTCTTAAGTGTGATTAAGGACCCACCACTCCAGT 2880
Db 2821 TTTCTGCTCAGCTTCTTAAGTGTGATTAAGGACCCACCACTCCAGT 2880
Qy 2881 AATTTTGTATTTTAAATGAGCGGGTTCGCCATTTGGCCAGGCTGTGTAATC 2940
Db 2881 AATTTTGTATTTTAAATGAGCGGGTTCGCCATTTGGCCAGGCTGTGTAATC 2940
Qy 2941 CTGACGTAAGTATGCTGCTGCTTGTGTTCCCAATTAAGGATGAACCACTGACCCA 3000
Db 2941 CTGACGTAAGTATGCTGCTGCTTGTGTTCCCAATTAAGGATGAACCACTGACCCA 3000
Qy 3001 CTAATTAAGTATTAATGATGATTAAGCAATTAAGAGATTTTCAATTTTCAATGAT 3060
Db 3001 CTAATTAAGTATTAATGATGATTAAGCAATTAAGAGATTTTCAATTTTCAATGAT 3060
Qy 3061 TTTTCTCTCTGCAAAATGCTTAAGTCTTGTGTTTCCCTTTTGGGGCAAGACACT 3120
Db 3061 TTTTCTCTCTGCAAAATGCTTAAGTCTTGTGTTTCCCTTTTGGGGCAAGACACT 3120

QY	3121	CATTAAATATCTGACATTTTTCCTATACAAGAGTATATCAAGTGTGTCTATAGA	3180
Db	3121	CATTAAATATCTGACATTTTTCCTATACAAGAGTATATCAAGTGTGTCTATAGA	3180
QY	3181	ACTGCGTGATTCGATTTATGTGTTTCTGATTCGATCTGTGTCCTTCATCCTTGAC	3240
Db	3181	ACTGCGTGATTCGATTTATGTGTTTCTGATTCGATCTGTGTCCTTCATCCTTGAC	3240
QY	3241	TCCTTTGGTATTTCACTGAATTTTCAACATTTGTGACAGAAAGAAAAGTAGAGACTCAG	3300
Db	3241	TCCTTTGGTATTTCACTGAATTTTCAACATTTGTGACAGAAAGAAAAGTAGAGACTCAG	3300
QY	3301	GAAAAATTAATTAATAAAGAACGCGCTTTGGCGGCGCGCGAATTC	3345
Db	3301	GAAAAATTAATTAATAAAGAACGCGCTTTGGCGGCGCGCGAATTC	3345

RESULT 4				
LOCUS	AX410642			
DEFINITION	AX410642	3345 bp	DNA	linear
ACCESSION	Sequence	3289 from Patent WO0229103.		PAT 14-JUN-2002
VERSION	AX410642			
KEYWORDS	AX410642.1	GI:21443347		
SOURCE	.			
ORGANISM	human.			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
REFERENCE	1.			
AUTHORS.	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.			
TITLE	Gene expression profiles in liver cancer			
JOURNAL	Patent: WO 0229103-A 3289 11-APR-2002;			
	GENE LOGIC INC (US)			
FEATURES	Location/Qualifiers			
source	1..3345			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/note="EMBL/GenBank Accession No. U07969"			
BASE COUNT	951 a	746 c	720 g	928 t
ORIGIN				

Query Match	Similarity	Score	DB	Length
Best Local	100.0%	3345	6	3345
Matches	Conservative	0	Mismatches	0
		0	Indels	0
		0	Gaps	0

QY	421	ATCACCAATAGAAGTGAAGGACATTAACGACATGACCCACGTTTCTCCAGTCAAGTAC	480
Db	421	ATCACCAATAGAAGTGAAGGACATTAACGACATGACCCACGTTTCTCCAGTCAAGTAC	480
QY	481	GAAAGGCTCAGTAAGAGGCAAACTCCGGCCGAGAAAGGCCCTTGTGTATGTCAATGCCACA	540
Db	481	GAAAGGCTCAGTAAGAGGCAAACTCCGGCCGAGAAAGGCCCTTGTGTATGTCAATGCCACA	540
QY	541	GACCTGGATGATCCGGCCACTCCCAATGGCCAGCTTTATTAACGAGTTGTCAACAGCTT	600
Db	541	GACCTGGATGATCCGGCCACTCCCAATGGCCAGCTTTATTAACGAGTTGTCAACAGCTT	600
QY	601	CCCATGATCAACAATGTCATGTAATTCATGATCAACAACAAACGGAGGCATCTCTT	660
Db	601	CCCATGATCAACAATGTCATGTAATTCATGATCAACAACAAACGGAGGCATCTCTT	660
QY	661	ACCCGAGAAGGATCTCGGAATTTGAATCCGTCGTAAGAAATCCCTTCCTTAATCTGAGATC	720
Db	661	ACCCGAGAAGGATCTCGGAATTTGAATCCGTCGTAAGAAATCCCTTCCTTAATCTGAGATC	720
QY	721	TCAGTGAAGGACATGGAGAGCCAGAGTGAATAATCCCTTCAGTATACACATCTGTGAT	780
Db	721	TCAGTGAAGGACATGGAGAGCCAGAGTGAATAATCCCTTCAGTATACACATCTGTGAT	780
QY	781	ATCATAGTGA CAGGAATATTGGAAGACCAAACCTGTGAGATGTGGAATACTCA	840
Db	781	ATCATAGTGA CAGGAATATTGGAAGACCAAACCTGTGAGATGTGGAATACTCA	840
QY	841	ACTGATCTCTCAACCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGGTGCAATAT	900
Db	841	ACTGATCTCTCAACCCATCAAAATCACTCAGGTGCGGTGGAATGATCCCGGTGCAATAT	900
QY	901	TCCCTAGTTGACAAAGAGAGTGGCCAAAGTCCCAATTTCAATGACAGAAAGAGAT	960
Db	901	TCCCTAGTTGACAAAGAGAGTGGCCAAAGTCCCAATTTCAATGACAGAAAGAGAT	960
QY	961	AATTAACGTGA CTGACCCCTTGAACCGAAGAAAGATGCATATGTTTTTTATGCAATT	1020
Db	961	AATTAACGTGA CTGACCCCTTGAACCGAAGAAAGATGCATATGTTTTTTATGCAATT	1020
QY	1021	GCAAAGATGATGACGAAACCACTTTCAATCCGCTGGAATTCATGTGAAAAGTTAAA	1080
Db	1021	GCAAAGATGATGACGAAACCACTTTCAATCCGCTGGAATTCATGTGAAAAGTTAAA	1080
QY	1081	GATATTATATATATATCAACCTACATGTCGTCACAGTAAACGTAATTTGAGGTCAGAG	1140
Db	1081	GATATTATATATATATCAACCTACATGTCGTCACAGTAAACGTAATTTGAGGTCAGAG	1140
QY	1141	AATGAACGACTGGGTAAACGATATCGGGACCTTACATGACATGACGGGATGAAGAAAT	1200
Db	1141	AATGAACGACTGGGTAAACGATATCGGGACCTTACATGACATGACGGGATGAAGAAAT	1200
QY	1201	ACTGCCAACAGTTTCTTAAACTTACAGAGATGTGTGACAAATCCCAAACCTTCCCATGAT	1260
Db	1201	ACTGCCAACAGTTTCTTAAACTTACAGAGATGTGTGACAAATCCCAAACCTTCCCATGAT	1260
QY	1261	GGACTCTTCTATCAAAACCTATGCTGAAATGTTACAGTTAAGCTTAAACAGTCTTTGAAG	1320
Db	1261	GGACTCTTCTATCAAAACCTATGCTGAAATGTTACAGTTAAGCTTAAACAGTCTTTGAAG	1320
QY	1321	AACCAAGATATCTCTCACTAACAATTAACGATGAGGTGTCAACAAAGATTTCAAGACC	1380
Db	1321	AACCAAGATATCTCTCACTAACAATTAACGATGAGGTGTCAACAAAGATTTCAAGACC	1380
QY	1381	CTTTGTTTGTGCAAAATCAACGTTATGATATCATGATCAGATCCCAATCTTTGAAAAA	1440
Db	1381	CTTTGTTTGTGCAAAATCAACGTTATGATATCATGATCAGATCCCAATCTTTGAAAAA	1440
QY	1441	TCAGATTATGGAACCTGACTCTGTGTCGAAGACAAACAATGGGTGTCACCACTTAAC	1500
Db	1441	TCAGATTATGGAACCTGACTCTGTGTCGAAGACAAACAATGGGTGTCACCACTTAAC	1500
QY	1501	ATCCAGGCACTGATGCTGATGAGCCATTTAATCGGAGTTCTTAATTCGTATCATATC	1560

Db 1501 ATCCAGCCCTGATGCTGATGACCCATTACTGCGGATTCTAATAATTCGTATCATATC 1560
 Qy 1561 ATAAAGGAGACAGTGAAGGAGCGCTGCGGGGTTGACACAGATCCCATACCAACCGGA 1620
 Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGCGGGGTTGACACAGATCCCATACCAACCGGA 1620
 Qy 1621 TATGCTAATTAATAAAGCTCTTGATTTGAAAACAGACAGCTGTTCCAACTGTGTC 1680
 Db 1621 TATGCTAATTAATAAAGCTCTTGATTTGAAAACAGACAGCTGTTCCAACTGTGTC 1680
 Qy 1681 AAAGCAGAAAATCCAGAGCTCTAGTGTGTGAGTAAAGTAAAGTAAAGTAAAGTAAAG 1740
 Db 1681 AAAGCAGAAAATCCAGAGCTCTAGTGTGTGAGTAAAGTAAAGTAAAGTAAAGTAAAG 1740
 Qy 1741 AAGTTCACGCTTAATGTAAGATGTAATGAATGAACACCTCAATTTCCCAACAGTATTC 1800
 Db 1741 AAGTTCACGCTTAATGTAAGATGTAATGAATGAACACCTCAATTTCCCAACAGTATTC 1800
 Qy 1801 CAAGCGAAAGTCAAGTGAAGATGTAATGAAGCACTAAAGTGAAGTGAAGTGAAGTGAAG 1860
 Db 1801 CAAGCGAAAGTCAAGTGAAGATGTAATGAAGCACTAAAGTGAAGTGAAGTGAAGTGAAG 1860
 Qy 1861 GATCCAGAAAGTCAATTAAGTATTAATCACTGAGGAGGAGACAGAAAGTGAAGTGAAGTGAAG 1920
 Db 1861 GATCCAGAAAGTCAATTAAGTATTAATCACTGAGGAGGAGACAGAAAGTGAAGTGAAGTGAAG 1920
 Qy 1921 ATTGACCAAGTCAATTAAGTATTAATCACTGAGGAGGAGACAGAAAGTGAAGTGAAGTGAAG 1980
 Db 1921 ATTGACCAAGTCAATTAAGTATTAATCACTGAGGAGGAGACAGAAAGTGAAGTGAAGTGAAG 1980
 Qy 1981 CCATATGAGGATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040
 Db 1981 CCATATGAGGATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040
 Qy 2041 GAGTTCCACCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 Db 2041 GAGTTCCACCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
 Qy 2101 ACGGCGTGTGTTCTTCTGATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGGCTACT 2160
 Db 2101 ACGGCGTGTGTTCTTCTGATCCCTCAGTGCACCTGGAAGTCTCATTTTGAAGGCTACT 2160
 Qy 2161 GATGATGATCAGCACTTAATTTGGGGGTTCCCATTTTACATTTTCCCTGCGAGTGAAGC 2220
 Db 2161 GATGATGATCAGCACTTAATTTGGGGGTTCCCATTTTACATTTTCCCTGCGAGTGAAGC 2220
 Qy 2221 TTACAAAACAGCTGGAAGTTCACAAAATCAATGTAATCAATGTAATGTAATGTAATGTAATGTAAT 2280
 Db 2221 TTACAAAACAGCTGGAAGTTCACAAAATCAATGTAATCAATGTAATGTAATGTAATGTAATGTAAT 2280
 Qy 2281 CACACAGACTTGAAGAGAGGCGTATGCTGTTGATCCGATCAATGATGAGGAGTGG 2340
 Db 2281 CACACAGACTTGAAGAGAGGCGTATGCTGTTGATCCGATCAATGATGAGGAGTGG 2340
 Qy 2341 CCACCTTTGGAAGGATGTTCTTACCACTTACATTTCTGAGTGTGTGTAAGGAGT 2400
 Db 2341 CCACCTTTGGAAGGATGTTCTTACCACTTACATTTCTGAGTGTGTGTAAGGAGT 2400
 Qy 2401 TGTTCGCGGAGAGGATCAGCAGCTGGAATACCACTGAGGAGTGGAGTGTATTA 2460
 Db 2401 TGTTCGCGGAGAGGATCAGCAGCTGGAATACCACTGAGGAGTGGAGTGTATTA 2460
 Qy 2461 CTGCTGACCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
 Db 2461 CTGCTGACCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
 Qy 2521 AAGGATTAAGGAGATTAATGTTGAAGTGTCAAGATCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2580
 Db 2521 AAGGATTAAGGAGATTAATGTTGAAGTGTCAAGATCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2580
 Qy 2581 AGCTGAATTTGAAAAGATGTTGAATTTATATGAAGTGTCAATTTTCAAGTGAAGTGAAGTGAAGTGAAG 2640
 Db 2581 AGCTGAATTTGAAAAGATGTTGAATTTATATGAAGTGTCAATTTTCAAGTGAAGTGAAGTGAAGTGAAG 2640

Db 2581 AGCTGAATTTGAAAAGATGTTGAATTTATATGAAGTGTCAATTTTCAAGTGAAGTGAAGTGAAGTGAAG 2640
 Qy 2641 TCTCATCCATTAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Db 2641 TCTCATCCATTAATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Qy 2701 TGTCTTATTAATTTGCTAATAATTTCTTTTGAAGTGAAGTGTGCTGTGCGCCAG 2760
 Db 2701 TGTCTTATTAATTTGCTAATAATTTCTTTTGAAGTGAAGTGTGCTGTGCGCCAG 2760
 Qy 2761 GCTGAGTACAGTGTGATGATCCAGCTCACTGCAACCTCCGCTCTGAGTCAATGA 2820
 Db 2761 GCTGAGTACAGTGTGATGATCCAGCTCACTGCAACCTCCGCTCTGAGTCAATGA 2820
 Qy 2821 TTTCTGCGCTCAGCTTCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2880
 Db 2821 TTTCTGCGCTCAGCTTCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2880
 Qy 2881 AATTTTGTATTTTATTAAGAGAGGAGGTTTCCCAATTTGCGCAGGCTGTGTAATC 2940
 Db 2881 AATTTTGTATTTTATTAAGAGAGGAGGTTTCCCAATTTGCGCAGGCTGTGTAATC 2940
 Qy 2941 CTGAGCTCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Db 2941 CTGAGCTCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Qy 3001 CCACTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
 Db 3001 CCACTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
 Qy 3061 TTTTCTCTCTGCAAAATGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3120
 Db 3061 TTTTCTCTCTGCAAAATGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3120
 Qy 3121 CATTAATAATTTCTGATCAATTTTCTTTATCAAGAGATTAATCAATGATGATGATGATGATGATGATGAT 3180
 Db 3121 CATTAATAATTTCTGATCAATTTTCTTTATCAAGAGATTAATCAATGATGATGATGATGATGATGATGAT 3180
 Qy 3181 ACTGCTGATTTGATTAATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
 Db 3181 ACTGCTGATTTGATTAATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
 Qy 3241 TCCCTTGTATTTCACTGATTTCAAAATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
 Db 3241 TCCCTTGTATTTCACTGATTTCAAAATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
 Qy 3301 GAAAAATTAATAATAAAGAGACAGCTTTGCGCGCGAATTC 3345
 Db 3301 GAAAAATTAATAATAAAGAGACAGCTTTGCGCGCGAATTC 3345
 RESULT 5
 HSDU07969
 LOCUS
 DEFINITION
 Human intestinal peptide-associated transporter HPT-1 mRNA,
 complete cds.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 3345)
 Dantzig, A.H., Hoskins, J., Tabae, L.B., Bright, S., Shepard, R.L.,
 Jenkins, I.L., Duckworth, D.C., Sportsman, R., Mackensen, D.,
 Roestek, P.R. and Skatrud, P.L.
 Association of intestinal peptide transport with a protein related
 to the cadherin superfamily
 Science 264 (5157), 430-433 (1994)
 JOURNAL
 MEDLINE
 PUBMED
 94204643
 8153632
 2 (bases 1 to 3345)

AUTHORS

Rosteck, P. R.
Direct Submission
Submitted (24-MAR-1994) Paul R. Rosteck Jr., Lilly Research
Laboratories, Eli Lilly and Co., Lilly Corporate Center,
Indianapolis, IN 46285, USA

FEATURES

Location/Qualifiers

1..3345

/organism="Homo sapiens"

/db_xref="taxon:9606"

/cell_line="Caco-2"

/cell_type="adenocarcinoma"

/tissue_type="colon"

88..2586

/codon_start=1

/product="intestinal peptide-associated transporter HPT-1"

/protein_id="AA19021.1"

/db_xref="GI:483392"

/translation="MILQAHLSLCILMLYLATGQGKFSGLPKMPFSIEQGE
SOIIPKPNPAPVTEFETGETDNIPIREBGLLYNRALDRETRSHNIOVALDAN
GIIVEGVPITIEVKDINDRPTLOSKEYSVONRPPKPLFYNAITLDLPATN
GOLYOIVIOLEPMNNMYFOINNKGTGISTREBSGELNPARNPSYNIIVSYDMG
OSNSPSTSVIIVITENTMKA PKVEMENSTDPPICTITVRMNDPEAGSIVDK
EKIPRPFSIDQSDIYITOPLDREKDAYFAVADETGKPLSTPLETHVKQIN
DNPPICSPVTFEVDENERLNSIGLTAHDEBENTANSFLNRYLVEQTPPLMDG
LFLIQYAGMLQAKOSLKKQDTPQVNLTFEVDKDFKTLCPQVINVIDINDOIPIE
KSDYGNLTLEADNITGSLTILIOATDEBPTSSKLIYH IKGDSBGRGVTDPH
NTGVILIKKPLDEPETAIVSNIVFAENPEPLVGVKNASPAKFTLLIVDNEAOF
SOHVPKYSVEDVALIGTKGNMTAKDEGLDISYLRGDRGLKIDHTVGEI.FSVAP
LDREAGSPTRVQVAVTEVGSSISVSERFLIMDVNDNPKLAKDTYGF.FCHPLSA
PGLIFPATDDHLEFRPHPTFLSGLSLONDMEVSKINGTARSTRTDEEAY
VILIRINDGRPLBGLIVSLPVTFCSCEFRPAHQGTITVGVAAGIILTLTLLV
IGIILAVFRIKDKGKNVESAOSEVSPPLRS"

CDS

BASE COUNT 951 a 746 c 720 g 928 t

Query Match 100.0%; Score 3345; DB 9; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GAATTCGCTCCACCTGAATGGAAGAAAGACTTTTAAACCACTTTTGTGACTTA 60
1 GAATTCGCTCCACCTGAATGGAAGAAAGACTTTTAAACCACTTTTGTGACTTA 60
61 CAGAAAGAAATTTGAATTAAGAAAATCTATGATCTTCAGGCCCATCTTCCCTGTG 120
61 CAGAAAGAAATTTGAATTAAGAAAATCTATGATCTTCAGGCCCATCTTCCCTGTG 120
121 CTTCTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAAGTTTATGAGACCCCTG 180
121 CTTCTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAAGTTTATGAGACCCCTG 180
181 AAACCCATGACATTTTCTATTTATGAAGGCCAAGACCGAGTCAATTAATTCAGATT 240
181 AAACCCATGACATTTTCTATTTATGAAGGCCAAGACCGAGTCAATTAATTCAGATT 240
241 AAGGCCAATCTCTGCTGTGAATTTTGAAGGCCAAGACCGAGTCAATTAATTCAGATT 240
241 AAGGCCAATCTCTGCTGTGAATTTTGAAGGCCAAGACCGAGTCAATTAATTCAGATT 240
301 ATGAAACGGAGGAGCTTCTGTATTAACAAGAGCTTGGACAGGGAACAGATCTACT 360
301 ATGAAACGGAGGAGCTTCTGTATTAACAAGAGCTTGGACAGGGAACAGATCTACT 360
361 CAAATCTCCAGGTTGAGCCCTGAGCGCTTAATGGAATTAATGAGGGGTCCAGTCCCT 420
361 CAAATCTCCAGGTTGAGCCCTGAGCGCTTAATGGAATTAATGAGGGGTCCAGTCCCT 420
421 ATCAACATGAAATGGAAGACATCAAGCAATCGAACCGCTTCTCCAGTCAAAATAC 480
421 ATCAACATGAAATGGAAGACATCAAGCAATCGAACCGCTTCTCCAGTCAAAATAC 480
481 GAAAGCTCAGTAAAGCAAACTCTGCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACA 540
481 GAAAGCTCAGTAAAGCAAACTCTGCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACA 540

481 GAAAGCTCAGTAAAGCAAACTCTGCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACA 540
541 GAAAGCTCAGTAAAGCAAACTCTGCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACA 540
541 GAAAGCTCAGTAAAGCAAACTCTGCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACA 540
541 GAAAGCTCAGTAAAGCAAACTCTGCGCCAGAAAGCCCTTCTTGTATGTCAATGCCACA 540
601 CCCATGATCAACATGATCATGATCTTCAAGATCAACAAACGAGGAGCATCTCTT 660
601 CCCATGATCAACATGATCATGATCTTCAAGATCAACAAACGAGGAGCATCTCTT 660
661 ACCCGAGAGGAGATCTCAGAAATGAAATCTGCTTAAGAAATCTTCTTAATTCGTGATC 720
661 ACCCGAGAGGAGATCTCAGAAATGAAATCTGCTTAAGAAATCTTCTTAATTCGTGATC 720
721 TCAGTAAAGCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
721 TCAGTAAAGCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
781 ATCATAGTGAACAGAAATATTTGAAAGCAACAAACCTGTGAGATGAGTGAAGAACTCA 840
781 ATCATAGTGAACAGAAATATTTGAAAGCAACAAACCTGTGAGATGAGTGAAGAACTCA 840
841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGACAAATAT 900
841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGACAAATAT 900
901 TCCCTTGTGACAAAGAAAGCTGCGCAAGATTTCCCAATTTTCAATTTGACCAAGAGAT 960
901 TCCCTTGTGACAAAGAAAGCTGCGCAAGATTTCCCAATTTTCAATTTGACCAAGAGAT 960
961 ATTTAGTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
961 ATTTAGTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
1021 GCAAGAGATGAGTGAAGAAACCACTTCATATCCGCTGAAATTTATGTAAGTAAAT 1080
1021 GCAAGAGATGAGTGAAGAAACCACTTCATATCCGCTGAAATTTATGTAAGTAAAT 1080
1081 GATATTAATGATTAATCCACTATGATCCGCTGACCAAGTAAATTTGATGAGGTCGAGAG 1140
1081 GATATTAATGATTAATCCACTATGATCCGCTGACCAAGTAAATTTGATGAGGTCGAGAG 1140
1141 AATGAAGCACTGGGTGAACGATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1141 AATGAAGCACTGGGTGAACGATTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
1201 ACTGCAACAGATTTCTTAACATCAAGAGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
1201 ACTGCAACAGATTTCTTAACATCAAGAGATTTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
1261 GGAATCTTCTTAATCCAACTATGCTGGAAGTGAACAGTGAAGTGAAGTGAAGTGAAG 1320
1261 GGAATCTTCTTAATCCAACTATGCTGGAAGTGAACAGTGAAGTGAAGTGAAGTGAAG 1320
1321 AAGCAAGATATCTCTGATTAACAACTTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
1321 AAGCAAGATATCTCTGATTAACAACTTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
1381 CTTGTTTGTGCAATCAACGTTATGATTAATCAATGATCAATGATCAATGATCAATGATCA 1440
1381 CTTGTTTGTGCAATCAACGTTATGATTAATCAATGATCAATGATCAATGATCAATGATCA 1440
1441 TCAGATTAAGAAACCTGACTCTTGTGTAAGACAAACATTTGGTCCACATCTTAAC 1500
1441 TCAGATTAAGAAACCTGACTCTTGTGTAAGACAAACATTTGGTCCACATCTTAAC 1500
1501 ATCCAGGAGCAGTATGATGATGAGCATTATCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
1501 ATCCAGGAGCAGTATGATGATGAGCATTATCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
1561 ATTAAGGAGGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
1561 ATTAAGGAGGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

QY 8 GTCTGACCACTGATGGAAGAAAAGACCTTTAACAACATTTTGTGACTTACAGAAG 67
DB 13 GTCTGACCACTGATGGAAGAAAAGACCTTTAACAACATTTTGTGACTTACAGAAG 72
QY 68 GAATTTGAATTAAGAAAATATGATTACTTACAGGCCCATCTTCACTCCCTGTGTCTTTA 127
DB 73 GAATTTGAATTAAGAAAATATGATTACTTACAGGCCCATCTTCACTCCCTGTGTCTTTA 132
QY 128 TGCTTTATTTGGCAACTGATATGAGCCAGAGGGGAAAGTTTAGTGGACCCCTGAACCA 187
DB 133 TGCTTTATTTGGCAACTGATATGAGCCAGAGGGGAAAGTTTAGTGGACCCCTGAACCA 192
QY 188 TGAATTTTCTATTTATTAAGAGCCCAAGACCGAGTCAATTAATTAATTCAGTTTAAGSCA 247
DB 193 TGACATTTTCTATTTATTAAGAGCCCAAGACCGAGTCAATTAATTAATTCAGTTTAAGSCA 252
QY 248 ATCTCTCTGTGTGACTTTTGAATTAATCTGGGAGACAGACACATATTTGTGATGAAC 307
DB 253 ATCTCTCTGTGTGACTTTTGAATTAATCTGGGAGACAGACACATATTTGTGATGAAC 312
QY 308 GGGAGGGAGCTTCTGTATTAACAAGAGCTTTGACAGGSAACAAGATCTACATCAATC 367
DB 313 GGGAGGGAGCTTCTGTATTAACAAGAGCTTTGACAGGSAACAAGATCTACATCAATC 372
QY 368 TCCAGGTTGAGCCCTGAGAGCTTAATGAAATTAATGAGAGGTCCAGTCCCTATCAACA 427
DB 373 TCCAGGTTGAGCCCTGAGAGCTTAATGAAATTAATGAGAGGTCCAGTCCCTATCAACA 432
QY 428 TAGAATGGAAGACATCAAGACATGACACCCAGTTCCTCAGTCAAGTCAAGAGCT 487
DB 433 TAAAGTGAAGGACATCAAGACATGACACCCAGTTCCTCAGTCAAGTCAAGAGCT 492
QY 488 CAGTAAGGCGAATCTCTGCGCCAGGAAAGCCCTTCTGTATGTCAAAGCCAGACCTGG 547
DB 493 CAGTAAGGCGAATCTCTGCGCCAGGAAAGCCCTTCTGTATGTCAAAGCCAGACCTGG 552
QY 548 ATGATCCGAGCACTCCCAATGAGCCTTTATTAACAGATTTGTCATCCAGTTCCTCAATGA 607
DB 553 ATGATCCGAGCACTCCCAATGAGCCTTTATTAACAGATTTGTCATCCAGTTCCTCAATGA 612
QY 608 TCAACAATGTCATGATCTTTCAGATCAACAACAACGAGAGCCATCTCTTAACCCGAG 667
DB 613 TCAACAATGTCATGATCTTTCAGATCAACAACAACGAGAGCCATCTCTTAACCCGAG 672
QY 668 AGGATATCCAGAAATTTGAATTCCTGTCAAGATCTTCTTAATCTGTGATCTCAATGA 727
DB 673 AGGATATCCAGAAATTTGAATTCCTGTCAAGATCTTCTTAATCTGTGATCTCAATGA 732
QY 728 AGGACATGGAGGCGCAGAGTGAAGATCTTTCAGTGAATCAACATCTGTGATATCAAG 787
DB 733 AGGACATGGAGGCGCAGAGTGAAGATCTTTCAGTGAATCAACATCTGTGATATCAAG 792
QY 788 TGAACAGAAATTTTGGAAAGACCAAAACCTGTGAGATGTGGAAAACCTCAACTGATC 847
DB 793 TGAACAGAAATTTTGGAAAGACCAAAACCTGTGAGATGTGGAAAACCTCAACTGATC 852
QY 848 CTACCCCATCAAAATCACTCAGGTGAGTGGAGATGATCCCGGTGCACAATATTCCTTAG 907
DB 853 CTACCCCATCAAAATCACTCAGGTGAGTGGAGATGATCCCGGTGCACAATATTCCTTAG 912
QY 908 TTGACAAAGAGAGCTGCAAGATTCCTCATTTTCAATTGACAGAGAGAGATATTTAGC 967
DB 913 TTGACAAAGAGAGCTGCAAGATTCCTCATTTTCAATTGACAGAGAGAGATATTTAGC 972
QY 968 TGACTAGCCCTTTGAGCCGAGAGAAAAGATGATATGTTTTTTATGACAGTTGCAAGG 1027
DB 973 TGACTAGCCCTTTGAGCCGAGAGAAAAGATGATATGTTTTTTATGACAGTTGCAAGG 1032
QY 1028 ATGAGTACGGAACCACTTTTCATATCCGCTGGAATTTCAATGTAAGTTAAAGTAAATTA 1087
DB 1033 ATGAGTACGGAACCACTTTTCATATCCGCTGGAATTTCAATGTAAGTTAAAGTAAATTA 1092

QY 1088 ATGATATTCACCTACATGTCGGTCAACAGTAAACCGTATTTGAGTCCAGAGATGTAAC 1147
DB 1093 ATGATATTCACCTACATGTCGGTCAACAGTAAACCGTATTTGAGTCCAGAGATGTAAC 1152
QY 1148 GACTGGGTAAACGATATGGGACCTTTATCTGACATGACAGAGATGAAGAAATATCTCCA 1207
DB 1153 GACTGGGTAAACGATATGGGACCTTTATCTGACATGACAGAGATGAAGAAATATCTCCA 1212
QY 1208 ACAGTTTTCTAAACTCAGAGATTGTGGAGCAACCTCCAAACTCCCATGGATGGAATCT 1267
DB 1213 ACAGTTTTCTAAACTCAGAGATTGTGGAGCAACCTCCAAACTCCCATGGATGGAATCT 1272
QY 1268 TCCATATCCAAACCTATGCTGGAATGTTTACAGTTAGCTTAAACAGTCCCTGAAGAGCAG 1327
DB 1273 TCCATATCCAAACCTATGCTGGAATGTTTACAGTTAGCTTAAACAGTCCCTGAAGAGCAG 1332
QY 1328 ATACTCTCTGATCAACTTAACGATAGAGGTGTGCAAAAGATTTCAAGACCTTTGTT 1387
DB 1333 ATACTCTCTGATCAACTTAACGATAGAGGTGTGCAAAAGATTTCAAGACCTTTGTT 1392
QY 1388 TTGTGCAAAATCAAGTTATTTGATATCAATGATCCCATCTTTGAAAATACAGAT 1447
DB 1393 TTGTGCAAAATCAAGTTATTTGATATCAATGATCCCATCTTTGAAAATACAGAT 1452
QY 1448 ATGAAAACCTGACTCTTGTGTAAGACACAACAATTGGGTCCACCATCTTAACATCCAGG 1507
DB 1453 ATGAAAACCTGACTCTTGTGTAAGACACAACAATTGGGTCCACCATCTTAACATCCAGG 1512
QY 1508 CCACTGATGCTGATGAGCCATTTACTGGGAGTTCTAAATTTCTGTATCATATCAATAAAG 1567
DB 1513 CCACTGATGCTGATGAGCCATTTACTGGGAGTTCTAAATTTCTGTATCATATCAATAAAG 1572
QY 1568 GAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATCAACACCGGATATGTCA 1627
DB 1573 GAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATCAACACCGGATATGTCA 1632
QY 1628 TAATTAATAAGCCTCTTGAATTTTGAACAAGACACCTGTTTCCAACTTTGTGTAAGAGCAG 1687
DB 1633 TAATTAATAAGCCTCTTGAATTTTGAACAAGACACCTGTTTCCAACTTTGTGTAAGAGCAG 1692
QY 1688 AAAATCCTGAGCCTCTAGTGTGTTGTGTGTAAGTACATGCAATGCTTTTGGCAAGTTCA 1747
DB 1693 AAAATCCTGAGCCTCTAGTGTGTTGTGTGTAAGTACATGCAATGCTTTTGGCAAGTTCA 1752
QY 1748 CGCTTAATGTGAAGAATGTGAATGAAGACCTCAATTTTCCCAACAGTATTCACAGCA 1807
DB 1753 CGCTTAATGTGAAGAATGTGAATGAAGACCTCAATTTTCCCAACAGTATTCACAGCA 1812
QY 1808 AAGTCAGTGAAGATGATGATATGAGCACTTAAAGTGGCAATGATGATCCCAAGATCCAG 1867
DB 1813 AAGTCAGTGAAGATGATGATATGAGCACTTAAAGTGGCAATGATGATGATCCCAAGATCCAG 1872
QY 1868 AAGTCTGACATTAAGCTATTCACAGGAGGACACAAGAGGTTGGCTTAAATTTGACC 1927
DB 1873 AAGTCTGACATTAAGCTATTCACAGGAGGAGACACAAGAGGTTGGCTTAAATTTGACC 1932
QY 1928 ACGTGACTGTGAGATCTTAAAGTGGCTCCCATTTGACAGAGAAGCCGGAAGTCCATATC 1987
DB 1933 ACGTGACTGTGAGATCTTAAAGTGGCTCCCATTTGAGAGAGAAGCCGGAAGTCCATATC 1992
QY 1988 GGGTACAAATGTGAGGCAAGAGATGAGGGGGTCTCTTAAAGCTCTGTGTCAAGATTC 2047
DB 1993 GGGTACAAATGTGAGGCAAGAGATGAGGGGGTCTCTTAAAGCTCTGTGTCAAGATTC 2052
QY 2048 ACCTGATCTTATGATGATGATGACAAACCTTCCAGGCTTACAGAGACTACAGGGCT 2107
DB 2053 ACCTGATCTTATGATGATGATGACAAACCTTCCAGGCTTACAGAGACTACAGGGCT 2112
QY 2108 TGTTCCTTGGCATCCCTCAGTGCACCTGGAAGTCAATTTTCCAGGCTACATGATGATG 2167
DB 2113 TGTTCCTTGGCATCCCTCAGTGCACCTGGAAGTCAATTTTCCAGGCTACATGATGATG 2172
QY 2168 ATCAGCACTTAATTTGGGGTCCCATTTTACATTTTCCCTCGGCAGTGAAGCTTACAA 2227

Db	2173	ATCAGCACTTATTTCCGGGGTCCCATTTTACATTTTCCCTCGSAGTGGAAAGCTTACAA	2232
QY	2228	ACGACGTGGAAAGTTTCCAAATTCATGCTACATGCCCCGACTGTCTACGACACAG	2287
Db	2233	ACGACGTGGAAAGTTTCCAAATTCATGCTACATGCCCCGACTGTCTACGACGACAG	2292
QY	2288	ACTTTGAGAGAGGGCGGTATGTGCTCTTGAATCCGCATCAATGATGGGGGTCCGACACT	2347
Db	2293	AGTTTGGAGAGAGGAGATATGTGCTCTTGAATCCGCATCAATGATGGGGGTCCGACACT	2352
QY	2348	TGAAAGCATGTGTTCTTACCACTTCAATCTGCAAGTTGTGTGGAAGAAAGTTGTTCC	2407
Db	2353	TGAAAGCATGTGTTCTTACCACTTCAATCTGCAAGTTGTGTGGAAGAAAGTTGTTCC	2412
QY	2408	GGCCAGCAGGTCCACCAACTGGAAATCCCACTGTGGCAGTGGCAGTTGTATCTGCTGA	2467
Db	2413	GGCCAGCAGGTCCACCAACTGGAAATCCCACTGTGGCAGTGGCAGTTGTATCTGCTGA	2472
QY	2468	CCACCCCTTCTGTGTAATGTGTAATTTTTAGCAGTTGTGTTATTCGGATTAAGAAAGATA	2527
Db	2473	CCACCCCTTCTGTGTAATGTGTAATTTTTAGCAGTTGTGTTATTCGGATTAAGAAAGATA	2532
QY	2528	AAGCAAAAGATATGTGTAAGAGTGTCAAGCATCTGAAGTCAAACTCTGGAAGCTGAA	2587
Db	2533	AAGCAAAAGATATGTGTAAGAGTGTCTAAGCATCTGAAGTCAAACTCTGGAAGCTGAA	2592
QY	2588	TTTGGAAAAGATGTTTGAATTTATATATAGCAAGTGTATTTTCAGCAACAACATCTCATC	2647
Db	2593	TTTGGAAAAGATGTTTGAATTTATATATAGCAAGTGTATTTTCAGCAACAACATCTCATC	2652
QY	2648	CTATTTACTTTTCATCTAACTGTCATATATATTTTTTAAACAGATATTCCTCTGTGCTT	2707
Db	2653	CTATTTACTTTTCATCTAACTGTCATATATATTTTTTAAACAGATATTCCTCTGTGCTT	2712
QY	2708	TAATATTTGCTAAATATTTCTTTTTTGAGGTGGAGTCTGTGCTGTGGCCAGGCTGGAG	2767
Db	2713	TAATATTTGCTAAATATTTCTTTTTTGAGGTGGAGTCTGTGCTGTGGCCAGGCTGGAG	2772
QY	2768	TACAGTGTGTGATCCAGACTCACTGCAACCTCGGCTCTGGGTTCACATGATTCCTCT	2827
Db	2773	TACAGTGTGTGATCCAGACTCACTGCAACCTCGGCTCTGGGTTCACATGATTCCTCTCT	2832
QY	2828	GCTCTAGCTCTTCAATGATGCTGGGTTTACAGGCAACCACCACTGCCAGCTAATTTTT	2887
Db	2833	GCTCTAGCTCTTCAATGATGCTGGGTTTACAGGCAACCACCACTGCCAGCTAATTTTT	2892
QY	2888	GTAATTTTTTAATAGACGGGGGTTTCGCAATTTGGCAGGCTGGTCTTGAATCTCGACGT	2947
Db	2893	GTAATTTTTTAATAGACGGGGGTTTCGCAATTTGGCAGGCTGGTCTTGAATCTCGACGT	2952
QY	2948	CAAGTGATCTGCTGCTGTGGTCTCCCAATACAGGCATGAACCACTGACCCCACTACTT	3007
Db	2953	CAAGTGATCTGCTGCTGTGGTCTCCCAATACAGGCATGAACCACTGACCCCACTACTT	3012
QY	3008	AGATATTTTCATGTGCTATATAGCATTAGAGAGATTTTTCATTTTTCATGACATTTTCTCT	3067
Db	3013	AGATATTTTCATGTGCTATATAGCATTAGAGAGATTTTTCATTTTTCATGACATTTTCTCT	3072
QY	3068	CTTCGCAATGCGCTAGCTACTGTGTGTTTTCCCTTTGGGGCAAGACAGCTCATTTAA	3127
Db	3073	CTTCGCAATGCGCTAGCTACTGTGTGTTTTCCCTTTGGGGCAAGACAGCTCATTTAA	3132
QY	3128	TATTCGTACATTTTTCTTTATCAAGAGATATATCAGTGTGTGCTCATAGAACTGCTCT	3187
Db	3133	TATTCGTACATTTTTCTTTATCAAGAGATATATCAGTGTGTGCTCATAGAACTGCTCT	3192
QY	3188	GGATTCATTTATGTTTTTTCTGATTCATCTGTGTCCTCTCAATCCTTGACTCTTTG	3247
Db	3193	GGATTCATTTATGTTTTTTCTGATTCATCTGTGTCCTCTCAATCCTTGACTCTTTG	3252
QY	3248	GTATTTCACTGAATTTCAACATTTGTGACAGAGAGAAAAAGTGAAGACTAGAAAAAT	3307

Db	3253	GIATTTGACGGAATTTCAAACTTTGTGCAGAGAGAAAAAAGCTGAGAGCTCAGAGAAAAAT	3313
OY	3308	AAATTAATTAAGAAGACGCTTTTGC	3333
Db	3313	AAATTAATTAAGAAGACGCTTTTCC	3338
RESULT 7			
AX335495	AX335495	3654 bp	DNA
LOCUS	Sequence 6004 from Patent WO0194629.		linear
DEFINITION	AX335495		PAT 09-JAN-2002
ACCESSION	AX335495		
VERSION	AX335495.1	GI:18126214	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,		
TITLE	Hortigan, S., Soppet, D. R. and Weaver, Z.		
JOURNAL	Cancer gene determination and therapeutic screening using signature gene sets		
FEATURES	Patent: WO 0194629-A 6004 13-DEC-2001;		
FEATURES	Avalon Pharmaceuticals (US)		
SOURCE	Location/Qualifiers		
BASE COUNT	1. 3654		
ORIGIN	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
Query Match	1052 a	804 c	777 g 1021 t
Best Local Similarity	99.1%;	Score 3314.8;	DB 6;
Matches 3319;	Conservative	0;	Mismatches 7;
			Indels 0;
			Gaps 0;
OY	8	GTCGCGACCACTGAATGGAAGAAAAGAGCTTTTAACACACATTTTGTGACTTACAGAAAG	67
Db	13	GTCGCGACCACTGAATGGAAGAAAAGAGCTTTTAACACACATTTTGTGACTTACAGAAAG	72
OY	68	GAATTTGAATTAAGAAAACCTATGATATCTTCAGGCCATCTTCACTCCCTGTGCTTCTTA	127
Db	73	GAATTTGAATTAAGAAAACCTATGATATCTTCAGGCCATCTTCACTCCCTGTGCTTCTTA	132
OY	128	TGCTTTATTTGGCACTGATATGGCCAAAGGGGAAAGTTTAGTGACCCCTGAAAACCA	187
Db	133	TGCTTTATTTGGCACTGATATGGCCAAAGGGGAAAGTTTAGTGACCCCTGAAAACCA	192
OY	188	TGACATTTTCATTTATGAAGGCCAACAACGAGTCAATTAATTCAGTTTAAGGCA	247
Db	193	TGACATTTTCATTTATGAAGGCCAACAACGAGTCAATTAATTCAGTTTAAGGCA	252
OY	248	ATCCTCTGCTGTGACCTTTGAATTAATCTGGGAGACAGACATAATTTGTGATAGAAC	307
Db	253	ATCCTCTGCTGTGACCTTTGAATTAATCTGGGAGACAGACATAATTTGTGATAGAAC	312
OY	308	GGGAGGAGCTTCTGTATTTACAACAGAGCCTTTGACAGGAGAAACAATCTACTACATC	367
Db	313	GGGAGGAGCTTCTGTATTTACAACAGAGCCTTTGACAGGAGAAACAATCTACTACATC	372
OY	368	TCCAGGTTTGACAGCCCTGAGAGCTTAATGGAATTTAAGTGGAGGGTCCAGTCCCTATCCCA	427
Db	373	TCCAGGTTTGACAGCCCTGAGAGCTTAATGGAATTTAAGTGGAGGGTCCAGTCCCTATCCCA	432
OY	428	TAGAAGTGAAGGACATCAACGACATGAGCCACGTTTCTCCAGTCAAAGTACGAAGCT	487
Db	433	TAGAAGTGAAGGACATCAACGACATGAGCCACGTTTCTCCAGTCAAAGTACGAAGCT	492
OY	488	CAGTAAAGCAGAACTTCGCCAGAGAAAGCCCTTCTTGATATGCAATGCCACAGACCTGG	547
Db	493	CAGTAAAGCAGAACTTCGCCAGAGAAAGCCCTTCTTGATATGCAATGCCACAGACCTGG	552
OY	548	ATGATCCGGGACCTCCCAATGGCCAGCTTTATTCAGAGATTTGATCCAGTCCCATGA	607

Db 553 ATGATCGGCGACATCCCAATGGCCAGCTTTATTAACGATGTCATCAGCTTCCACATGA 612
Qy TCAACAAATGTCATGTACTTTGAGATCAACAACAAAAGGGGAGCCATCTTATCCGAG 667
Db 608 TCAACAAATGTCATGTACTTTGAGATCAACAACAAAAGGGGAGCCATCTTATCCGAG 672
Qy 613 TCAACAAATGTCATGTACTTTGAGATCAACAACAAAAGGGGAGCCATCTTATCCGAG 672
Qy 668 AGGAGATCTCAGAAATTAATCTGCTAAGATCTTCTTAATCTGAGTCTCAGTGA 727
Db 673 AGGAGATCTCAGAAATTAATCTGCTAAGATCTTCTTAATCTGAGTCTCAGTGA 732
Qy 728 AGGAGATCTCAGAAATTAATCTGCTAAGATCTTCTTAATCTGAGTCTCAGTGA 787
Db 733 AGGAGATCTCAGAAATTAATCTGCTAAGATCTTCTTAATCTGAGTCTCAGTGA 792
Qy 788 TGAAGAGATTAATTTGAAAAGCACAAACCTGAGAGATGAGTGAAGAACTCAATCATC 847
Db 793 TGAAGAGATTAATTTGAAAAGCACAAACCTGAGAGATGAGTGAAGAACTCAATCATC 852
Qy 848 CTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAAAATATTCCTAG 907
Db 853 CTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAAAATATTCCTAG 912
Qy 908 TTGACAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTTGACAGAGAGAGATATTTAGC 967
Db 913 TTGACAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTTGACAGAGAGAGATATTTAGC 972
Qy 968 TGAAGAGATTAATTTGAAAAGCACAAACCTGAGAGATGAGTGAAGAACTCAATCATC 1027
Db 973 TGAAGAGATTAATTTGAAAAGCACAAACCTGAGAGATGAGTGAAGAACTCAATCATC 1032
Qy 1028 ATGAGTACGGAAGAACCTTTCATATCCGCTGGAATTTCAATTTGAAAGTAAAGATATTA 1087
Db 1033 ATGAGTACGGAAGAACCTTTCATATCCGCTGGAATTTCAATTTGAAAGTAAAGATATTA 1092
Qy 1088 ATGATATCACTCACTCACTGTCCTGACAGATTAACCGTATTTGAGTTCAGAGAGATGAC 1147
Db 1093 ATGATATCACTCACTCACTGTCCTGACAGATTAACCGTATTTGAGTTCAGAGAGATGAC 1152
Qy 1148 GATCGGATACAGATGAGGAGACCTTCTGACATGACAGAGATGAGAGAAATATCTGCCA 1207
Db 1153 GATCGGATACAGATGAGGAGACCTTCTGACATGACAGAGATGAGAGAAATATCTGCCA 1212
Qy 1208 ACAGTTTCTTAACTACAGAGATGAGGAGAACTCCCAAACTTCCATGATGAGTCT 1267
Db 1213 ACAGTTTCTTAACTACAGAGATGAGGAGAACTCCCAAACTTCCATGATGAGTCT 1272
Qy 1268 TCCATATCCAAACCTATGCTGGAATGTTACAGTTAGCTTAAACAGTCTTGAAGAGCAG 1327
Db 1273 TCCATATCCAAACCTATGCTGGAATGTTACAGTTAGCTTAAACAGTCTTGAAGAGCAG 1332
Qy 1328 ATACTCTGATCACTTAATGATGAGGATGCTGACAAAGATTTCAACACCTTTGTT 1387
Db 1333 ATACTCTGATCACTTAATGATGAGGATGCTGACAAAGATTTCAACACCTTTGTT 1392
Qy 1388 TTGTCGAAATCAAGTATATGATATCAATGATCAGATCCCATCTTTGAAATATCAGATT 1447
Db 1393 TTGTCGAAATCAAGTATATGATATCAATGATCAGATCCCATCTTTGAAATATCAGATT 1452
Qy 1448 ATGGAACCTGACTCTTGTGGAAGACAACAACTTGGGTCCACATCTTAAACCATCCAG 1507
Db 1453 ATGGAACCTGACTCTTGTGGAAGACAACAACTTGGGTCCACATCTTAAACCATCCAG 1512
Qy 1508 CCACTGATGCTGATGAGGACCTTATCTGGAGTCTTAAATCTGATCATATCATTAAG 1567
Db 1513 CCACTGATGCTGATGAGGACCTTATCTGGAGTCTTAAATCTGATCATATCATTAAG 1572
Qy 1568 GAGACAGTGAAGGAGCCCTGGGGTTGACACAGATCCCATCAACACCGAGATATGTC 1627
Db 1573 GAGACAGTGAAGGAGCCCTGGGGTTGACACAGATCCCATCAACACCGAGATATGTC 1632
Qy 1628 TAATTAAGAGCTCTTGTATTTGAAAACAGAGCTGTTTCAACATTTGTTCAAGAG 1687

Db 1633 TAATTAAGAGCTCTTGTATTTGAAAACAGAGCTGTTTCAACATTTGTTCAAGAG 1692
Qy 1688 AAATCCGAGCCCTAGTGTGTTGTTGTAAGTACATGCAAGTCTTTGCCAAGTCA 1747
Db 1693 AAATCCGAGCCCTAGTGTGTTGTTGTAAGTACATGCAAGTCTTTGCCAAGTCA 1752
Qy 1748 CGCTTATTTGACAGATGTAATGAAAGACACCTCAATTTTCCCAACGATATTCAGAGCA 1807
Db 1753 CGCTTATTTGACAGATGTAATGAAAGACACCTCAATTTTCCCAACGATATTCAGAGCA 1812
Qy 1808 AAGTCAGTGAAGATGTAAGTATAGGCACTTAAAGTGGCAATGATCTGCCAAGATCCAG 1867
Db 1813 AAGTCAGTGAAGATGTAAGTATAGGCACTTAAAGTGGCAATGATCTGCCAAGATCCAG 1872
Qy 1868 AAGTCAGTGAAGATGTAAGTATAGGCACTTAAAGTGGCAATGATCTGCCAAGATCCAG 1927
Db 1873 AAGTCAGTGAAGATGTAAGTATAGGCACTTAAAGTGGCAATGATCTGCCAAGATCCAG 1932
Qy 1928 ACGTGAAGTGAAGATCTTATGATGAGTCTCATTTGACAGAGAGCGGAAGTCCATATC 1987
Db 1933 ACGTGAAGTGAAGATCTTATGATGAGTCTCATTTGACAGAGAGCGGAAGTCCATATC 1992
Qy 1988 GGGTACAGAGTGGGCGCACAGAGTAAAGGGGGTCTTCTTAAGCTCTGTCTCAGAGTCC 2047
Db 1993 GGGTACAGAGTGGGCGCACAGAGTAAAGGGGGTCTTCTTAAGCTCTGTCTCAGAGTCC 2052
Qy 2048 ACGTGAAGTGAAGATGTAAGTACACCTCCGAGGCTAGCCAGAGATCAACGGGCT 2107
Db 2053 ACGTGAAGTGAAGATGTAAGTACACCTCCGAGGCTAGCCAGAGATCAACGGGCT 2112
Qy 2108 TGTTCCTTCCCATCCCTCAGTGCACCTGGAAGTCTCATTTTCCAGAGGCTATGATGAT 2167
Db 2113 TGTTCCTTCCCATCCCTCAGTGCACCTGGAAGTCTCATTTTCCAGAGGCTATGATGAT 2172
Qy 2168 ATCAGCACTTATTTGGGGTCCCATTTTATCATTTTCCCTGGCAGTGAAGTTCACAA 2227
Db 2173 ATCAGCACTTATTTGGGGTCCCATTTTATCATTTTCCCTGGCAGTGAAGTTCACAA 2232
Qy 2228 ACGACGAGGAGATTTCAAAATCAATGATATCAATCCGACGTCATCAACGAGCACAG 2287
Db 2233 ACGACGAGGAGATTTCAAAATCAATGATATCAATCCGACGTCATCAACGAGCACAG 2292
Qy 2288 ACTTTGAGAGAGAGGCGTATGTCGTTGATCCGATCAATGATGAGGGTCCGCCACCT 2347
Db 2293 ACTTTGAGAGAGAGGCGTATGTCGTTGATCCGATCAATGATGAGGGTCCGCCACCT 2352
Qy 2348 TGAAGGCAATGTTTCTTACAGATTAATTTGCAAGTGTGTGGAAGAGTGTTC 2407
Db 2353 TGAAGGCAATGTTTCTTACAGATTAATTTGCAAGTGTGTGGAAGAGTGTTC 2412
Qy 2408 GGCAGAGAGTACACGAGCTGGGATCCACGTCGTGGCATGCGATTTGATATCTGCTGA 2467
Db 2413 GGCAGAGAGTACACGAGCTGGGATCCACGTCGTGGCATGCGATTTGATATCTGCTGA 2472
Qy 2468 CCAACCTTCTGATGATGATATTTTGAAGTGTGTGTTATCCCATTAAGAGATTA 2527
Db 2473 CCAACCTTCTGATGATGATATTTTGAAGTGTGTGTTATCCCATTAAGAGATTA 2532
Qy 2528 AAGGCAAGATTAATTTGAAGTGTCAAGCATCTGAAGTCAACCTCTGAGAGCTGA 2587
Db 2533 AAGGCAAGATTAATTTGAAGTGTCAAGCATCTGAAGTCAACCTCTGAGAGCTGA 2592
Qy 2588 TTTGAAAAGAGATGTTGAATTTATATGCAAGTGTATTTCAAGCAACCATCTCATC 2647
Db 2593 TTTGAAAAGAGATGTTGAATTTATATGCAAGTGTATTTCAAGCAACCATCTCATC 2652
Qy 2648 CTATTAATTTTCAATCAAGTCAATTAATTTTAAACAGATATTTCCCTCTGCTCT 2707
Db 2653 CTATTAATTTTCAATCAAGTCAATTAATTTTAAACAGATATTTCCCTCTGCTCT 2712
Qy 2708 TAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTCTCTGCTGCGCACAGCTGAG 2767
Db 2713 TAATATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTCTCTGCTGCGCACAGCTGAG 2772

QY 1148 GACTGGTAACAGTATGGGACCCTTACTGACATGACAGGGATGANGAANAATCTGCCA 1207
DB 1153 GACTGGGTAAACAGATGGGGACCCCTTACTGACATGACAGGGATGANGAANAATCTGCCA 1212
QY 1208 ACAGTTTCTAACTACAGAGTTGTGGAGCAAACTCCCAAACCTCCGATGGATGGACTCT 1267
DB 1213 ACAGTTTCTAACTACAGAGTTGTGGAGCAAACTCCCAAACCTCCGATGGATGGACTCT 1272
QY 1268 TCCATATCCAAACCTATGCTGGAATGTATACAGTTAGCTTAAACAGTCCCTGGAAGAAGCAG 1327
DB 1273 TCCATATCCAAACCTATGCTGGAATGTATACAGTTAGCTTAAACAGTCCCTGGAAGAAGCAG 1332
QY 1328 ATACTCTCAGTACATCTTAACATGAGAGTGTCTGACAAAGATTCAAGACCTTTGTT 1387
DB 1333 ATACTCTCAGTACATCTTAACATGAGAGTGTCTGACAAAGATTCAAGACCTTTGTT 1392
QY 1388 TTGTGCAATCAACGTTATGATATCAATGATCCCATCTTTGAAATAACATAT 1447
DB 1393 TTGTGCAATCAACGTTATGATATCAATGATCCCATCTTTGAAATAACATAT 1452
QY 1448 ATGGAACCTGACTCTTGTGAGACACAAACATTGGGTCACCATCTTAACATCCAGG 1507
DB 1453 ATGGAACCTGACTCTTGTGAGACACAAACATTGGGTCACCATCTTAACATCCAGG 1512
QY 1508 CCACTGATGCTGATGAGCCATTATCTGGAGTTCTAAATTTCTGTATCATATCAATAAGG 1567
DB 1513 CCACTGATGCTGATGAGCCATTATCTGGAGTTCTAAATTTCTGTATCATATCAATAAGG 1572
QY 1568 GAGACGCTGAGGAGCGCTGGGGTTGACACAGATCCCATACCAACACCGGATATGTCA 1627
DB 1573 GAGACGCTGAGGAGCGCTGGGGTTGACACAGATCCCATACCAACACCGGATATGTCA 1632
QY 1628 TAATTAATAAAGCCTCTGATTTTGAACAAGAGCTGTTTCCAACTTGTGTTCAAGCAG 1687
DB 1633 TAATTAATAAAGCCTCTGATTTTGAACAAGAGCTGTTTCCAACTTGTGTTCAAGCAG 1692
QY 1688 AAAATCTGAGCCTCTAGTGTGTGTGTGAAGTACAAATGCAAGTTCTTTTCCAGTTCA 1747
DB 1693 AAAATCTGAGCCTCTAGTGTGTGTGTGAAGTACAAATGCAAGTTCTTTTCCAGTTCA 1752
QY 1748 GGTCTATGAGACAGATGTAATGAACACCTCAATTTTCCCAACAGTATTCMAAGCA 1807
DB 1753 GGTCTATGAGACAGATGTAATGAACACCTCAATTTTCCCAACAGTATTCMAAGCA 1812
QY 1808 AAGTCACTGAGATGTAGTATAGGACCTAAAGGGCAATGTGACTGCCAAGGATCCAG 1867
DB 1813 AAGTCACTGAGATGTAGTATAGGACCTAAAGGGCAATGTGACTGCCAAGGATCCAG 1872
QY 1868 AAGTCTGACATTAAGCTATCTAGAGGGGAGACACAAAGGTTGGCTTAAATTTGACC 1927
DB 1873 AAGTCTGACATTAAGCTATCTAGAGGGGAGACACAAAGGTTGGCTTAAATTTGACC 1932
QY 1928 ACGTGACTGATGATTTTATGTGTGCTCAATTTGACAGAGAACCCGAAAGTCCATATC 1987
DB 1933 ACGTGACTGATGATTTTATGTGTGCTCAATTTGACAGAGAACCCGAAAGTCCATATC 1992
QY 1988 GGGTCAAGAGTGTGGGACACAGAGTAGGGGGTCTTCTTAAGTCTGTGTCAGAGTTCC 2047
DB 1993 GGGTCAAGAGTGTGGGACACAGAGTAGGGGGTCTTCTTAAGTCTGTGTCAGAGTTCC 2052
QY 2048 ACCTGATCTTATGATGTGATGACAAACCTCCAGAGCTAGCCAAAGACCTAGACGGGCT 2107
DB 2053 ACCTGATCTTATGATGTGATGACAAACCTCCAGAGCTAGCCAAAGACCTAGACGGGCT 2112
QY 2108 TGTTCCTTGTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTTCGAGGCTAGTATGATG 2167
DB 2113 TGTTCCTTGTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTTCGAGGCTAGTATGATG 2172
QY 2168 ATCAGCACTTATTTGGGGTCCCATTTTTCATTTTCCCTGGAGTGAAGCTTCAAA 2227
DB 2173 ATCAGCACTTATTTGGGGTCCCATTTTTCATTTTCCCTGGAGTGAAGCTTCAAA 2232
QY 2228 ACGACTGGGAAGTTTCCAAATCAATGTACTGATGCCGACTGTCTTACAGGACACAG 2287

DB 2233 ACGACTGGGAAGTTTCCAAATCAATGTACTGATGCCGACTGTCTTACAGGACACAG 2292
QY 2288 ACTTTGAGAGAGGGGATATGTGTCTGTGATTCGCAATCAATGATGGGGGTGGGCAACCT 2347
DB 2293 AGTTTGAAGAGAGGATATGTGTCTGTGATTCGCAATCAATGATGGGGGTGGGCAACCT 2352
QY 2348 TGAAGGCAATTTTCTTTTCAACAGTTACATCTGCAAGTTGTGGAAGAAAGTTGTTCC 2407
DB 2353 TGAAGGCAATTTTCTTTTCAACAGTTACATCTGCAAGTTGTGGAAGAAAGTTGTTCC 2412
QY 2408 GGCACAGAGGTCAACGACGTGGGATACCACTGTGGGCATGCGAGTTGTATCTGCTGA 2467
DB 2413 GGCACAGAGGTCAACGACGTGGGATACCACTGTGGGCATGCGAGTTGTATCTGCTGA 2472
QY 2468 CCAACCTTCTGGGATTTGTGATTAATTTTGAAGTGTGTTATCCGATTAAGAAGATA 2527
DB 2473 CCAACCTTCTGGGATTTGTGATTAATTTTGAAGTGTGTTATCCGATTAAGAAGATA 2532
QY 2528 AAGCAAAAGATTAATGTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGAAAGTGA 2587
DB 2533 AAGCAAAAGATTAATGTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGAAAGTGA 2592
QY 2588 TTTGAAAAGGATGTTGAAATTTTATATAGCAAGTGTATTTGACAAACAATCTCATC 2647
DB 2593 TTTGAAAAGGATGTTGAAATTTTATATAGCAAGTGTATTTGACAAACAATCTCATC 2652
QY 2648 CTATTACTTTTCACTCAAGTGTATTAATTTTAAACAGATTTTCCCTGTTCCT 2707
DB 2653 CTATTACTTTTCACTCAAGTGTATTAATTTTAAACAGATTTTCCCTGTTCCT 2712
QY 2708 TAATATTTGCTAAATTTTCTTTTGAAGTGAAGTCTGTGCGCCAGGCTGAG 2767
DB 2713 TAATATTTGCTAAATTTTCTTTTGAAGTGAAGTCTGTGCGCCAGGCTGAG 2772
QY 2768 TACAGTGTGTGATCCCAAGTCACTGCAACTTCGCTCTGAGTTTCAATGATCTCT 2827
DB 2773 TACAGTGTGTGATCCCAAGTCACTGCAACTTCGCTCTGAGTTTCAATGATCTCT 2832
QY 2828 GCTCAGCTTCCCAATGATGAGTGTGTTTACAGGACCCACCAATGCCAGCTAATTTT 2887
DB 2833 GCTCAGCTTCCCAATGATGAGTGTGTTTACAGGACCCACCAATGCCAGCTAATTTT 2892
QY 2888 GTATTTTATTAAGACGGGGTTCGCAATTTGGCCAGGCTGGTCTTGAACCTCTGACGT 2947
DB 2893 GTATTTTATTAAGACGGGGTTCGCAATTTGGCCAGGCTGGTCTTGAACCTCTGACGT 2952
QY 2948 CAAATGATCTGCTGCTTGTGCTCCAAATACAGGATGAACCACTGCAACCTTACTT 3007
DB 2953 CAAATGATCTGCTGCTTGTGCTCCAAATACAGGATGAACCACTGCAACCTTACTT 3012
QY 3008 AGATATTTCACTGTGCTATAGACATTAAGAGATTTTCAATTTTCCATGACATTTTCCCT 3067
DB 3013 AGATATTTCACTGTGCTATAGACATTAAGAGATTTTCAATTTTCCATGACATTTTCCCT 3072
QY 3068 CTCTGCAATAGGCTTACACTTGTGTCTTCCCTTTGGGGCAACAGACTTAA 3127
DB 3073 CTCTGCAATAGGCTTACACTTGTGTCTTCCCTTTGGGGCAACAGACTTAA 3132
QY 3128 TATCTGATCAATTTTCTTTTATCAAGAGATATACAGTGTGCTCATAGAACTGCT 3187
DB 3133 TATCTGATCAATTTTCTTTTATCAAGAGATATACAGTGTGCTCATAGAACTGCT 3192
QY 3188 GGAATTCATTTATGTTTCTTATGATTCATCTGTGTCCCTTATCTTGACTCTTTG 3247
DB 3193 GGAATTCATTTATGTTTCTTATGATTCATCTGTGTCCCTTATCTTGACTCTTTG 3252
QY 3248 GTATTCACATGATTTCAAACTTGTGAGAGAAAGAAAGAGAGACCTGAGAAAT 3307
DB 3253 GTATTCACATGATTTCAAACTTGTGAGAGAAAGAAAGAGAGACCTGAGAAAT 3312
QY 3308 AAATAAATAAAGAACAGCCTTTTGC 3333

KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 1 (bases 1 to 2499)
 REFERENCES
 Danczig, A.H., Hoskins, J.A., and Skatrud, P.L.
 TITLE Mammalian influx peptide transporter
 JOURNAL Patent: US 562085-A 2 15-APR-1997;
 FEATURES
 Location/Qualifiers
 source 1..2499
 /organism="unknown"
 BASE COUNT 728 a 559 c 570 g 642 t
 ORIGIN
 Query Match 74.7%; Score 2499; DB 6; Length 2499;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

88 ATGATATTGAGGCGCCATCTTCACTCCCTGCTGCTTCTTATGCTTATTTGGCAACTGGA 147
 1 ATGATATTGAGGCGCCATCTTCACTCCCTGCTGCTTCTTATGCTTATTTGGCAACTGGA 60
 148 TATGCGCAAGAGGGAGGAGTTAGTGAACCCCTGAAACCATGACATTTTCTATTTATGAA 207
 61 TATGCGCAAGAGGGAGGAGTTAGTGAACCCCTGAAACCATGACATTTTCTATTTATGAA 120
 208 GGGCCAGAACCGAGTCAATTAATTCAGTTTAAAGGCCAATCCTCTGCTGATCTTTT 267
 121 GGGCCAGAACCGAGTCAATTAATTCAGTTTAAAGGCCAATCCTCTGCTGATCTTTT 180
 268 GAATTAATCTGGGAGAGACAGACATATTTGTATAGAACGGGGAGGAGCTTGTATTAAC 327
 181 GAATTAATCTGGGAGAGACAGACATATTTGTATAGAACGGGGAGGAGCTTGTATTAAC 240
 328 AACAGAGCTTGGACAGGGAAACAAGATCTTACTCAATCTCCAGTTTGCAGCCCTGAC 387
 241 AACAGAGCTTGGACAGGGAAACAAGATCTTACTCAATCTCCAGTTTGCAGCCCTGAC 300
 388 GCTAATGAGATTAATAGTGAAGGGTCCAGTCCCTATCAACCATTAAGATGAAGACATCAAC 447
 301 GCTAATGAGATTAATAGTGAAGGGTCCAGTCCCTATCAACCATTAAGATGAAGACATCAAC 360
 448 GACAAATGACCCAGCTTCTCCAGTCAAGTACGAAGGCTCAGTAAGGACAGACTCTGAC 507
 361 GACAAATGACCCAGCTTCTCCAGTCAAGTACGAAGGCTCAGTAAGGACAGACTCTGAC 420
 508 CCAGAAAGAGCCCTTCTGTATGTCAATGCCACAGACTGTATGATCCGGCCACTCCCAAT 567
 421 CCAGAAAGAGCCCTTCTGTATGTCAATGCCACAGACTGTATGATCCGGCCACTCCCAAT 480.
 568 GGGCAGCTTATTAACCAATGTGTCAAGCTTCCCATGATCAACAATGTCAATGTACTTT 627
 481 GGGCAGCTTATTAACCAATGTGTCAAGCTTCCCATGATCAACAATGTCAATGTACTTT 540
 628 CAGATCAACAACAACGGGAGCCATCTCTTAACCGGAGGGAGTCCAGAAATTGAT 687
 541 CAGATCAACAACAACGGGAGCCATCTCTTAACCGGAGGGAGTCCAGAAATTGAT 600
 688 CCTCTAAGATCTTCTTATATCTGTGATCTTCAAGTGAAGGACATGGAGGCCAGT 747
 601 CCTCTAAGATCTTCTTATATCTGTGATCTTCAAGTGAAGGACATGGAGGCCAGT 660
 748 GAGATTCCTTCAAGTATCAACATCTGTGATATCATATGACAGAGATATTTTGGAAA 807
 661 GAGATTCCTTCAAGTATCAACATCTGTGATATCATATGACAGAGATATTTTGGAAA 720
 808 GCAACCAAAACCTGTGAGATGTGAAAACCAATGATCTTCAACCCATCAAAATTCAT 867
 721 GCAACCAAAACCTGTGAGATGTGAAAACCAATGATCTTCAACCCATCAAAATTCAT 780
 868 CAGGTGCGGTGATATCCGGGTGACAAATTTCTTATGATGAACAAGAGAGCTGGCA 927
 781 CAGGTGCGGTGATATCCGGGTGACAAATTTCTTATGATGAACAAGAGAGCTGGCA 840

928 AGATCCCAATTTTCAATGACAGAGAGAGATATTTAGTACTCAGCCCTTGACCGA 987
 841 AGATCCCAATTTTCAATGACAGAGAGAGATATTTAGTACTCAGCCCTTGACCGA 900
 988 GAAGAAAGAGATGATATGTTTTTATGAGTTCAGAGATGATGACGAAACCACTT 1047
 901 GAAGAAAGAGATGATATGTTTTTATGAGTTCAGAGATGATGACGAAACCACTT 960
 1048 TCATATCCGTGGAATTTCAATGAAAGTTAAAGATTAATATATCACTACATGT 1107
 961 TCATATCCGTGGAATTTCAATGAAAGTTAAAGATTAATATATCACTACATGT 1020
 1108 CCGTCAACAGTAAACGATTTTGGTCCAGAGATGAACGATGGGTAAACATATCGGG 1167
 1021 CCGTCAACAGTAAACGATTTTGGTCCAGAGATGAACGATGGGTAAACATATCGGG 1080
 1168 ACCCTTACTGACATGACAGGATGAAGAAATATCTGCCAACGTTTCTTAACTACAGG 1227
 1081 ACCCTTACTGACATGACAGGATGAAGAAATATCTGCCAACGTTTCTTAACTACAGG 1140
 1228 ATTGTGAGCAAACTCCCAACTTCCCATGATGATCTTCTTAATCCAACTATGCT 1287
 1141 ATTGTGAGCAAACTCCCAACTTCCCATGATGATCTTCTTAATCCAACTATGCT 1200
 1288 GGAATGTTACGTTAGCTAAACAGTCTTGAAGAGAGATATCTCTCAGTACAACTTA 1347
 1201 GGAATGTTACGTTAGCTAAACAGTCTTGAAGAGAGATATCTCTCAGTACAACTTA 1260
 1348 ACGATAGAGGTGTCTGCAAAAGATTTCAAGACCCTTTGTTTGTGCAATCAACGTTATT 1407
 1261 ACGATAGAGGTGTCTGCAAAAGATTTCAAGACCCTTTGTTTGTGCAATCAACGTTATT 1320
 1408 GATATCAATGATGATATCCCATCTTTGAAATTCAGATTAATGAAACCTGATGCT 1467
 1321 GATATCAATGATGATATCCCATCTTTGAAATTCAGATTAATGAAACCTGATGCT 1380
 1468 GAAGACAAACAATGGGTCACCATCTTAACCATCAGGACATGATGCTGATGAGCA 1527
 1381 GAAGACAAACAATGGGTCACCATCTTAACCATCAGGACATGATGCTGATGAGCA 1440
 1528 TTTACTGGAGATTCTAATAATCTGTATCATATCAATAAAGGAGACAGTGAAGGACGCTG 1587
 1441 TTTACTGGAGATTCTAATAATCTGTATCATATCAATAAAGGAGACAGTGAAGGACGCTG 1500
 1588 GGGGTGACACAGATCCCATCAACAACCGGATATGTCTAATTAAGGCTCTGAT 1647
 1501 GGGGTGACACAGATCCCATCAACAACCGGATATGTCTAATTAAGGCTCTGAT 1560
 1648 TTTGAACACAGCTGTTTCAACATGTGTCAAGACAGAAATCCTGAGCCTCTAGT 1707
 1561 TTTGAACACAGCTGTTTCAACATGTGTGTCAAGACAGAAATCCTGAGCCTCTAGT 1620
 1708 TTTGTGTGAAGTACAAATGCAAGTTCTTTTGCAGAGTTCAAGCTTATGTGACAGATGTG 1767
 1621 TTTGTGTGAAGTACAAATGCAAGTTCTTTTGCAGAGTTCAAGCTTATGTGACAGATGTG 1680
 1768 AATGAACACTCAATTTTCCCAACAGTATTCGAAGCGAAAGTCAATGAGTATGCT 1827
 1681 AATGAACACTCAATTTTCCCAACAGTATTCGAAGCGAAAGTCAATGAGTATGCT 1740
 1828 AATGAACACTCAATGAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 1887
 1741 AATGAACACTCAATGAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 1800
 1888 TCACTGAGGGGAGACAAAGAGTTTGAATTAATGACCACTGATGATGATGATGAT 1947
 1801 TCACTGAGGGGAGACAAAGAGTTTGAATTAATGACCACTGATGATGATGATGATGAT 1860
 1948 AGTGTGCTCCATTTGACAGAGAAAGCGGAAAGTCAATATGAGGATCAAGTGTGCGACA 2007
 1861 AGTGTGCTCCATTTGACAGAGAAAGCGGAAAGTCAATATGAGGATCAAGTGTGCGACA 1920

QY	2008	GAAGTAGGGGGGCTTTCTCTTAAGCTCTGTGTCAAGATTCACACTGATCTTATGGATGTG	2067
Db	1921	GAAGTAGGGGGGCTTTCTCTTAAGCTCTGTGTCAAGATTCACACTGATCTTATGGATGTG	1980
QY	2068	AATGACAACCCCTCCAGGCTAGCCAAAGGACTACAGGGGCTGTCTTCCTGGCAATCCCTC	2127
Db	1981	AATGACAACCCCTCCAGGCTAGCCAAAGGACTACAGGGGCTGTCTTCCTGGCAATCCCTC	2040
QY	2128	AGTGCAACCTGGAAGTCTCATTTTTCGAGGCTACTGATGATACAGACTTTATTCGGGGT	2187
Db	2041	AGTGCAACCTGGAAGTCTCATTTTTCGAGGCTACTGATGATGATACAGACTTTATTCGGGGT	2100
QY	2188	CCCCATTTTACATTTTTCCTCGGCGAGTGGAACTTACAAACAGACTGGAAATTCCAA	2247
Db	2101	CCCCATTTTACATTTTTCCTCGGCGAGTGGAACTTACAAACAGACTGGAAATTCCAA	2160
QY	2248	ATCANGGTAATCATGATGCCGAGTGTCTTCCAGGCGACACAGACTTTGAGAGAAAGGGCGAT	2307
Db	2161	ATCANGGTAATCATGATGCCGAGTGTCTTCCAGGCGACACAGACTTTGAGAGAAAGGGCGAT	2220
QY	2308	GTGCTTTTGATCCGATCAATGATGAGGGGTGGCCACCCTTGAGAGGCAATGTTTCTTTA	2367
Db	2221	GTGCTTTTGATCCGATCAATGATGAGGGGTGGCCACCCTTGAGAGGCAATGTTTCTTTA	2280
QY	2368	CCAGTTAATCTCTGCAGTGTGTGTGGAAGAAAGTTGTTTCGGGCGACAGCTCACCAACT	2427
Db	2291	CCAGTTAATCTCTGCAGTGTGTGTGGAAGAAAGTTGTTTCGGGCGACAGCTCACCAACT	2340
QY	2428	GGAATACCACTGTGGGAGATGGCAGATTGATCTGCTGACCAACCCTCTGATGATGGT	2487
Db	2341	GGAATACCACTGTGGGAGATGGCAGATTGATCTGCTGACCAACCCTCTGATGATGGT	2400
QY	2488	ATAATTTTACGAGTTGTGTTTATCCGCAATAAAGAGATAAAGCAAGATATAGTTGAA	2547
Db	2401	ATAATTTTACGAGTTGTGTTTATCCGCAATAAAGAGATAAAGCAAGATATAGTTGAA	2460
QY	2548	AGTGCTCAAGCATCTGAAGTCAAACCTCTGAGAACTGCA	2586
Db	2461	AGTGCTCAAGCATCTGAAGTCAAACCTCTGAGAACTGCA	2499

	RESULT	..11			
	LOCUS	181197			
	DEFINITION	Sequence 2 from patent US 5710018.	2499 bp	DNA	linear
	ACCSSION	181197			PAT 10-JUN-1998
	VERSION	181197.1	GI:3209467		
	KEYWORDS				
	SOURCE	Unknown.			
	ORGANISM	Unknown.			
	REFERENCE	Unclassified. 1 (bases 1 to 2499)			
	AUTHORS	Dantzig,A.H., Hoskins,J Ann. and Skatrud,P.L.			
	TITLE	Mammalian influx peptide transporter			
	JOURNAL	Patent: US 5710018-A 2 20 -JAN-1998;			
	FEATURES	Location/Qualifiers 1..2499 source /organism="unknown"			
OY	BASE COUNT	728 a	559 c	570 g	642 t
OY	ORIGIN				
	Query Match	74.7% ; Score 2499; DB 6; Length 2499; Best Local Similarity 100.0%; Pred. No. 0; Matches 2499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	db				
	148	TATGCGCAAGGGGAAGTTTAACTGCACCCCTGAACCATTCATTCTTAATTATGAA	207		
	61	TATGGCCAAAGGGGAAGTTTAGTGACCCCTGAACCATTCATTCTTAATTATGAA	120		
OY	88	ATGATACTTCAGGCCAATCCTCACCTCCGTGTCTTCTTAATGCTTATTTGGCAACTGGA	147		
OY	Db	1 ATGATACTTCAGGCCCAATCCTCACCTCCGTGTCTTCTTAATGCTTATTTGGCAACTGGA	60		

QY	208	GGCCAAAGAA	CCGAGTCAAA	TTATATATTCAGTTTAA	GGCCAAATCTCTGCTGACTTTT	267	
Db	121	GGCCAAAGAA	CCGAGTCAAA	TTATATATTCAGTTTAA	GGCCAAATCTCTGCTGACTTTT	180	
QY	268	GAACTTAACT	GGGGAGACAGACA	CTATTTTGTGATGAA	CGGGAGGACTTCTGATTAC	327	
Db	181	GAACTTAACT	GGGGAGACAGACA	CTATTTTGTGATGAA	CGGGAGGACTTCTGATTAC	240	
QY	328	AACAGAGCCTT	GGACGAGGAA	CAAGATCTACTCA	CAATCTCAGGTTTACGCCCTGGAC	387	
Db	241	AACAGAGCCTT	GGACGAGGAA	CAAGATCTACTCA	CAATCTCAGGTTTACGCCCTGGAC	300	
QY	388	GCTAATGAA	ATTATATGAGGGG	GCAGTCCCTATCACC	ATTGAAGTGAAGACATCAAC	447	
Db	301	GCTAATGAA	ATTATATGAGGGG	GCAGTCCCTATCACC	ATTGAAGTGAAGACATCAAC	360	
QY	448	GACATTCGAC	CCACGTTTCTCCAGTCAAA	AGTTCGAAGCTCAGTAA	GGCAGAACTCTGCG	507	
Db	361	GACATTCGAC	CCACGTTTCTCCAGTCAAA	AGTTCGAAGCTCAGTAA	GGCAGAACTCTGCG	420	
QY	508	CAAGGAAAGC	CCCTTCTGTATGTCA	ATGTCACAGACTGATCA	CAATGTCAGTCACTCCCAAT	567	
Db	421	CAAGGAAAGC	CCCTTCTGTATGTCA	ATGTCACAGACTGATCA	CAATGTCAGTCACTCCCAAT	480	
QY	568	GGCCAGCTTT	TATTTATACAGATGTCAT	CCAGCTTCCCATGATCA	CAATGTCATGTACTTT	627	
Db	481	GGCCAGCTTT	TATTTATACAGATGTCAT	CCAGCTTCCCATGATCA	CAATGTCATGTACTTT	540	
QY	628	CAGATTCACA	CAACAAAACGGAGCCAT	CTCTCTTACCCGAGAGGAT	CTCAGGAATTGAAT	687	
Db	541	CAGATTCACA	CAACAAAACGGAGCCAT	CTCTCTTACCCGAGAGGAT	CTCAGGAATTGAAT	600	
QY	688	CTGTCTAAGA	ATTCCTTCCATATAT	TGTGATCTCAGTGAAGGA	CATGGGAGGCCAGAGT	747	
Db	601	CTGTCTAAGA	ATTCCTTCCATATAT	TGTGATCTCAGTGAAGGA	CATGGGAGGCCAGAGT	660	
QY	748	GAGAAATTC	CTTCAGTATACCA	CACTCTGTGATATCATAGTACAGAGAA	TATTTTGGAAA	807	
Db	661	GAGAAATTC	CTTCAGTATACCA	CACTCTGTGATATCATAGTACAGAGAA	TATTTTGGAAA	720	
QY	808	GCACTCAAAA	CTGTGAGATGTTGGAAA	CTCMACTGATCTC	CACCCATCAAAATCACT	867	
Db	721	GCACTCAAAA	CTGTGAGATGTTGGAAA	CTCMACTGATCTC	CACCCATCAAAATCACT	780	
QY	868	CAGGTGGG	GTGGAAATGATCCG	GGTGCAAAATATTCCTA	GTGTGCAAAAGAAAGGTGCA	927	
Db	781	CAGGTGGG	GTGGAAATGATCCG	GGTGCAAAATATTCCTA	GTGTGCAAAAGAAAGGTGCA	840	
QY	928	AGATTTCC	ATTTTCAATTTGAC	CCAGAGAGAAATTTTA	CGTGACTCAGCCCTTGGACCGA	987	
Db	841	AGATTTCC	ATTTTCAATTTGAC	CCAGAGAGAAATTTTA	CGTGACTCAGCCCTTGGACCGA	900	
QY	988	GAAAGAAAGAT	GCATATGTTTTTTTTT	TATGACAGTGC	AAAGATGAGTACGAAAAACACTT	1047	
Db	901	GAAAGAAAGAT	GCATATGTTTTTTTTT	TATGACAGTGC	AAAGATGAGTACGAAAAACACTT	960	
QY	1048	TCATATCCG	CTGAGAAATCATGTAA	AAATTAATAATTA	TGATAATCCACCTACAGT	1107	
Db	961	TCATATCCG	CTGAGAAATCATGTAA	AAATTAATAATTA	TGATAATCCACCTACAGT	1020	
QY	1108	CCGTCAACA	GTAAACCGTATTTTGA	AGGTCCAGAGATGAA	CACTGCGGTAA	CAGTATCGGG	1167
Db	1021	CCGTCAACA	GTAAACCGTATTTTGA	AGGTCCAGAGATGAA	CACTGCGGTAA	CAGTATCGGG	1080
QY	1168	ACCCTTAC	TGCACTGACATGAC	AGGATGAAAGAAATTA	CTGCAACAGTTTTCTTAA	CTACAGG	1227
Db	1081	ACCCTTAC	TGCACTGACATGAC	AGGATGAAAGAAATTA	CTGCAACAGTTTTCTTAA	CTACAGG	1140
QY	1228	ATTGTGAG	CAAACTCCCAAAC	CTTCCCATGGA	TGGAATCTTCCATATCCAA	ACCTATGCT	1287
Db	1141	ATTGTGAG	CAAACTCCCAAAC	CTTCCCATGGA	TGGAATCTTCCATATCCAA	ACCTATGCT	1200
QY	1288	GGAATGTT	ACAGTTAGCTTAA	CAGTCTCTTGAA	GAAGCAAGATATCTCTC	AGTACACTTA	1347

D	b	1201	GGAAATGTTACAGTTAGCTTAACAGAGCTCCTTGAGAGAGCAAGTAAGTCTCTCAAGTACAACTTA	1260
Q	y	1348	ACGATAGAGGTGTCTGACAAAGATTTCAGAGCCCTTTGTTTGTGCAAAATCAACGTTATT	1407
D	b	1261	ACGATAGAGGTGTCTGACAAAGATTTCAGAGCCCTTTGTTTGTGCAAAATCAACGTTATT	1320
Q	y	1408	GATATCAATGATCAAGATCCCATCTTTGAAAATAATCAATTAAGAAACCTGACTCTTGCT	1467
D	b	1321	GATATCAATGATCAAGATCCCATCTTTGAAAATAATCAATTAAGAAACCTGACTCTTGCT	1380
Q	y	1468	GAAGACACAAACATTTGGGTCACCATCTTAAACATCCAGGCCACTGATNGCTGATGAGCCA	1527
D	b	1381	GAAGACACAAACATTTGGGTCACCATCTTAAACATCCAGGCCACTGATNGCTGATGAGCCA	1440
Q	y	1528	TTTACTGGAGTTCTTAAATTTCTGTATCATATCATTAAGGGAGACAGTGAAGGACGCGCTG	1587
D	b	1441	TTTACTGGAGTTCTTAAATTTCTGTATCATATCATTAAGGGAGACAGTGAAGGACGCGCTG	1500
Q	y	1588	GGGGTTGACACAGATCCCATACCAACCGGATATGTCAATTAATAAAAGCCTCTTGAT	1647
D	b	1501	GGGGTTGACACAGATCCCATACCAACCGGATATGTCAATTAATAAAAGCCTCTTGAT	1560
Q	y	1648	TTTGAACACAGAGCTGTTCCAACTATGTGTTCAAAGAGAAATCCGAGCCTCTAGTG	1707
D	b	1561	TTTGAACACAGAGCTGTTCCAACTATGTGTTCAAAGAGAAATCCGAGCCTCTAGTG	1620
Q	y	1708	TTTGGTGTGAAGTACAAATGCAAGTCTTTTTCGCAAGTTCACGCTTAATGTGACAGATGTG	1767
D	b	1621	TTTGGTGTGAAGTACAAATGCAAGTCTTTTTCGCAAGTTCACGCTTAATGTGACAGATGTG	1680
Q	y	1768	AATGAAGACCTCAATTTTCCCAACAGTATTCGAAGGAAAGTCAAGTGAAGATGTAGCT	1827
D	b	1681	AATGAAGACCTCAATTTTCCCAACAGTATTCGAAGGAAAGTCAAGTGAAGATGTAGCT	1740
Q	y	1828	ATTAGGCACTAAGTGGGCAATGTACATGACCAGATCCAGAAAGTCTGGAATATAGCAT	1887
D	b	1741	ATTAGGCACTAAGTGGGCAATGTACATGACCAGATCCAGAAAGTCTGGAATATAGCAT	1800
Q	y	1888	TCACTGAAGGGGAGACACAGAAGTTGGCTTAAATTTGACACAGTGAAGTGAAGATCTTT	1947
D	b	1801	TCACTGAAGGGGAGACACAGAAGTTGGCTTAAATTTGACACAGTGAAGTGAAGATCTTT	1860
Q	y	1948	AGTGTGGCTCCATTGGAACAGAGAAAGCCGGAAGTCCATATCGGGTACAGATGTGGCCACA	2007
D	b	1861	AGTGTGGCTCCATTGGAACAGAGAAAGCCGGAAGTCCATATCGGGTACAGATGTGGCCACA	1920
Q	y	2008	GAAGTGAAGGGGGGTCTTCTTAAAGCTCTGTGTCAAGTTTCCACTGATCCTTATGAGATGTG	2067
D	b	1921	GAAGTGAAGGGGGGTCTTCTTAAAGCTCTGTGTCAAGTTTCCACTGATCCTTATGAGATGTG	1980
Q	y	2068	AATGACAAACCTCCAGGCTAGACCAAGAGACTACACGGGCTGTTCTTCTGCAATCCCTC	2127
D	b	1981	AATGACAAACCTCCAGGCTAGACCAAGAGACTACACGGGCTGTTCTTCTGCAATCCCTC	2040
Q	y	2128	AGTGCACCTGGAAGTCTCATTTTTCGAGGCTATCTGATGTATCAGACCTTAATTTTCGAGGT	2187
D	b	2041	AGTGCACCTGGAAGTCTCATTTTTCGAGGCTATCTGATGTATCAGACCTTAATTTTCGAGGT	2100
Q	y	2188	CCCCATTTTAATTTTCCCTCGGACGTGGAAGCTTACAAAACGACTGGAAGTTTCCAAA	2247
D	b	2101	CCCCATTTTAATTTTCCCTCGGACGTGGAAGCTTACAAAACGACTGGAAGTTTCCAAA	2160
Q	y	2248	ATCAATGATACATGATGCCGACGTGTCTACAGGCAACAGACTTTGAGAGAGGGCGTAT	2307
D	b	2161	ATCAATGATACATGATGCCGACGTGTCTACAGGCAACAGACTTTGAGAGAGGGCGTAT	2220
Q	y	2308	GTCGCTCTTGATCCGATCAATGATGTGGGGTGGCCACCTTGAAGGCAATGTTCTTTTA	2367
D	b	2221	GTCGCTCTTGATCCGATCAATGATGTGGGGTGGCCACCTTGAAGGCAATGTTCTTTTA	2280
Q	y	2368	CCAGTTACATTTCTGCAATTTGTGTGGAAGAAATTGTTTCGGCCACGAGGTCACAGACT	2427

Db	2281	CAGATTACATTCGCAAGTGTGTGGAAAGAAGTAGTTGTTCCGGCAGCAGAGCTCACCGACT	2344
Oy	2428	GGGATACCACCTGTGGGCATGCGAGTGTTGATATCTGTGACCAACCCCTCTGGGATTTGT	2487
Db	2341	GGGATVACCACCTGTGGGCATGCGAGTTGTTGATATCTGTGACCAACCCCTCTGGGATTTGT	2400
Oy	2488	ATAATTTTACAGTGTGTTTATCCGCATTAAGAAGATATAAAGCAAAGATTAATGTTGA	2547
Db	2401	ATAATTTTACAGTGTGTTTATCCGCATTAAGAAGATATAAAGCAAAGATTAATGTTGA	2460
Oy	2548	AGTGCTCAAGCATCTGAAGTCAAACTCTGAGAAGCTGA	2586
Db	2461	AGTGCTCAAGCATCTGAAGTCAAACTCTGAGAAGCTGA	2499
RESULT 12			
RNCADHER		3628 bp	mRNA linear ROD 13-SEP-1994
LOCUS	R.norvegicus (Sprague Dawley) mRNA for cadherin.		
DEFINITION	X78997		
ACCESSION	X78997.1 GI:505562		
VERSION	cadherin.		
KEYWORDS	Rattus norvegicus.		
SOURCE	Rattus norvegicus.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Rattus. 1 (bases 1 to 3628) Berndorff,D., Gessner,R., Kreft,B., Schnoy,N., Lajous-Petter,A.M., Loch,N., Reutter,W., Horstach,M. and Tauber,R. Liver-intestine cadherin: molecular cloning and characterization of a novel Ca(2+)-dependent cell adhesion molecule expressed in liver and intestine J. Cell Biol. 125 (6), 1353-1369 (1994)		
REFERENCE	1 (bases 1 to 3628)		
AUTHORS	Tauber,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
MEDLINE	and Biochemie, Universitaetsklinikum Rudolf Virchow, Freie Uni, Berlin, Spandauer Damm 130, 14050 Berlin, FRG		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	94266966		
AUTHORS	8207063		
TITLE	2 (bases 1 to 3628)		
JOURNAL	Tauber,R.		
MEDLINE	Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie		
PUBMED	942669		

polyA_signal 3605..3610
 BASE COUNT 1023 a 835 c 811 g 959 t
 ORIGIN
 Query Match 53.0%; Score 1774.4; DB 10; Length 3628;
 Best Local Similarity 80.4%; Pred. No. 0;
 Matches 2108; Conservative 0; Mismatches 501; Indels 13; Gaps 2;
 9 TCTCGACCACTGAATGGAAGAAAAGACTTTTAAACCACTTTTGTACTTACAGAAG 68
 247 TCTCCAGGCGTGAATGACGACCAAGCTTTTAAAGAGTATTTCTTACTTACGAAAG 306
 69 AATTGAATTAAGAAAATCTATGATCTTACAGCCCATCTTCACTCCCTGTCTTCTAT 128
 307 AGTCCAGATGAAGAGC-AACATGATCTCTGCCCACTTCACTTCTGTGTCTTCTAC 365
 129 GCTTATTTGGCACTGGATATGSCCAAGAGGGAATTTAGTGGACCCCTGAAACCAT 188
 366 GCTTATTTGACGTGGCCATATGSCCAAGAGGGAATTTGAGCGGCCCTGAAACCAT 425
 189 GACATTTTCTATTTATGAAAGCCCAAGACGATCAATTTATTTCCAGTTTAAAGCCAA 248
 426 GACATTTCTATTTTATTTGAAGCCCAAGACGATCAATTTATTTCCAGTTTAAAGCCAA 485
 249 TCTCTCTGCTGTACTTTTGAATTAAGTGGGAGACAGACAACTATTTGTATGTAAGC 308
 486 CCTCCAGCTGTGACTTTTGTAGCTTAAAGAGAGAGCGATGTATTTTAAATGAAAA 545
 309 GAGAGGACTTCTGATTTCAACAGAGCTTGGACAGGGAACAAAGATCTACTCAATCT 368
 546 GGAATGACTTCTTATATCAACAAGATCTGGAACAGAAAAACAAGGCGTTCACTACT 605
 369 CCAAGTTGACCCCTGAGCCTTAATGGAATTTATGTAAGGATCTCACTGCTTACCAT 428
 606 GCAAGCTTGACCTTGAATTTCTAGAGAGCTATAGTGAATGSCCACTCCCATCATAT 665
 429 AGAAGTGAAGACATCAACAACTGACCCAGCTTTCTCAAGTCAAAAGTACGAAGCTC 488
 666 AGAAGTGAAGACATCAACAACTGACCCAGCTTTCTCAAGTCAAAAGTACGAAGCTC 725
 489 AGTAAGCCAAATCTTGTGCCCCAGAAAGCCCTTCTGTATGTCAATTTCCACAGACTG 548
 726 AGTAAGCCAAATCTTGTGCCCCAGAAAGCCCTTCTGTATGTCAATTTCCACAGACTG 785
 549 TGATCCGCGCACTCCCAATGAGCTTTTATTAACAGATTTGTAATCCAGCTTCCCATAT 608
 786 TGATCCGCGCACTCCCAATGAGCTTTTATTAACAGATTTGTAATCCAGCTTCCCATAT 845
 609 CAACAATGTATGATCTTTCAGATCAACAAACGAGGACCACTCTCTTACCCGAGA 668
 846 CAACAATGTATGATCTTTCAGATCAACAAACGAGGACCACTCTCTTACCCGAGA 905
 669 GGGATCTCAGGAATTTGAATCTGTGAAGATCTTCTTAATCTGTGATCTCAAGTGA 728
 906 AGGATCTCAGGAATTTGAATCTTAAGAAATCTGTATCAACCTGTGTCTCTGTGA 965
 729 GGAATCTCAGGAATTTGAATCTTCAAGTATCAACCACTGTGTGATCTCAATGAT 788
 966 GGAATCTCAGGAATTTGAATCTTCAAGTATCAACCACTGTGTGATCTCAATGAT 1025
 789 GACAGAAATATTTGGAAGACCAACAAACCTGTGTGATGTGGAAGAACTCAACTGATC 848
 1026 CAGAGAAATATTTGGAAGACCAACAAACCTGTGTGATGTGGAAGAACTCAACTGATC 1085
 849 TCAACCCCATTAATTAATCTCAGGTGGGTGAATGATCCCGGTGCAATATTTCTTATG 908
 1086 TCAACCCCATTAATTAATCTCAGGTGGGTGAATGATCCCGGTGCAATATTTCTTATG 1145
 909 TGACAAAGAGAGCTGCAAGATTTCCATTTTCAATTTGACAGAGAGAGATTTTAAAGT 968
 1146 CAACAAGAGAGCTGCAAGATTTTCAATTTTCAATTTGACAGAGAGAGATTTTAAAGT 1205
 969 GACTGACCCCTTGGACCGAGAGAAAAAGATGCAATGTTTTTTTATGCAAGTTGCAAGAG 1028

1206 GACTGACCCCTTGGACCGTAGAGAAAAAGACTCTCATGTTTTCTTGTCCACTCCAGAGA 1265
 1029 TGAGTACGGAACCACTTTCATATCGCTGGAAATTCATGTAAGTTAAAGTAAATTTAA 1088
 1266 TGGAATGGAACCACTTTCATATCGCTGGAAATTCATGTAAGTTAAAGTAAATTTAA 1325
 1089 TGATTAATCCACTCATATGTCCTGACAGTAAACCTGATTTGAGAGTCCAGAGATGAACG 1148
 1326 CGAACACCCACCAATGTCCTCTCTCAAGTACCTGATTTGAGAGTCCAGAGATGAACG 1385
 1149 ACTGGGTAAACATATGAGGACCTTATCTGACATGAACAGGATGAAGAAAAATCTGCCA 1208
 1386 CTGGGTAGATGAGATTTGGATCTTTGTCTGCCACGATATGATGAAGCTAAATATCAA 1445
 1209 CAGTTTCTTAATTAAGATTTGTGAAGCAACTCCCAACTCCCATGATGATGATCTT 1268
 1446 CAGTATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
 1269 CTAATTCAAAACCTATGCTGAAATGTTACAGTAAAGCTAAACAGTCTTGAAGAGCAGA 1328
 1506 TCTCATTTGACGATATGAGAGAAAGTCCAGTTGGTAAAGCTCTTGAAGAGCAGA 1565
 1329 TACTCTCAGTACCACTTAAAGATAGAGGTGCTGAACAAAGATTTCAAGACCTTGT 1388
 1566 CAGTCTCAGTACCACTTAAAGATAGAGGTGCTGAACAAAGATTTCAAGACCTTGT 1625
 1389 TGTGCAATCAACGATTTATGATATCAATGATCAATGCCATTTTGAAGAAATCAATTA 1448
 1626 TCTTCAAGTCAAGTATGATATGATATGATATGATATGATATGATATGATATGAT 1685
 1449 TGAACAACTGACCTTCTGTAAGACCAACAAATGAGTCCACATCTTAAACATCCAGG 1508
 1686 TGAACAACTGACCTTCTGTAAGACCAACAAATGAGTCCACATCTTAAACATCCAGG 1745
 1509 CACTGATGCTGATGACCAATTAAGTGAAGTCTTAAATTTCTGATCATATCAATTAAG 1568
 1746 TACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
 1569 AGACAGTGAAGGACGCTGAGGCTTGAACACATGCCCATCAACACCGATATGATCAT 1628
 1806 AGACAGTGAAGGACGCTGAGGCTTGAACACATGCCCATCAACACCGATATGATCAT 1865
 1629 AATTAACAAAGCTCTTGAATTTGAACACAGAGTGTTCACATTTGTGTCAACAGAG 1688
 1866 GATCAACAAAGCTCTTGAATTTGAACACAGAGTGTTCACATTTGTGTCAACAGAG 1925
 1926 AAGCCTGAGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1985
 1749 GCTTATTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
 1986 GCTGACCGTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2045
 1809 AGTCAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1868
 2046 CTTTACGAGAGATCTGCTATGAGACTAAAGTGGACCTGATGATGATGATGATGATGAT 2105
 1869 AGTCTGACATTAAGTATTTCACTGAGGAGACACAAAGTGTGCTTAAATTTGACCA 1928
 2106 AGGCTGACATGATGATTTCACTGAGAGACAAATTAAGAGTGTGCTTAAATTTGACCA 2165
 1929 CGTGACTGTGATGATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1988
 2166 CGTGACCGGCGCATATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2225
 1989 GGTACAGTGTGAGGCGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2048
 2226 AGTACAGTGTGAGGCGACAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2285
 2049 CCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2108

Db	Accession	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	Source
Db	2286	CTGTGCTCCTATG	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2109	GTTCCTTCGCAT	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2346	GTTCTTCCTGCAT	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2169	TCACACCTTAT	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2406	CCAGACAGCA	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2229	CGACTGGGAA	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2466	TGACTGGGAA	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2289	CTTTGAGGAG	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2526	CTTTGAGGAA	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2349	GGAAGGCA	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2586	GGAAGGCA	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2409	GCAAGCAG	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2646	GCAAGCAG	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2469	CACCTTCG	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2706	CACATTCG	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2529	AGGCAAGT	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex 2133), Fax: 03-5285-1150							
Db	2756	--GGAAAT	mus musculus	D87912	1.1	GI:9392287	mus musculus	Mus musculus	Submitted (18-SEP-1996)	Kazuo Ohnishi	National Institute of Health, Department of Immunology, 1-23-1, Toyama, Shinjuku-ku, Tokyo 162, Japan (E-mail: ohnishi@kenhi.go.jp)	Tel: 03-5285-1111 (ex							

Query Match	Similarity	80.3%	Pred. No. 0	Mismatches	504	Indels	14	Gaps	3
Base Count	967 a	766 c	774 g	907 t					
Origin									
Query Match	52.6%	Score 1758.6	DB 10	Length 3414					
Best Local Similarity	80.3%	Pred. No. 0							
Matches 2105	Conservative	0	Mismatches	504	Indels	14	Gaps	3	
8	GTCTGACCACTGAATGGAAGAAAGACCTTTTAAACCACTTTTGACCTTAAGAAAG	67							
12	GTCTCCACGGCTGATGAGACGACAAAACCTTCAGACGA-TTTCTTACTTACTGAAAG	70							
68	GAATTTGAATPAAAGAAACATGATGATCTTCAGCGCCACTCTCCCTGTCTTCTTA	127							
71	GAGTCCAGTGAAGAGACAA-GATGGTGTCTGCCAGCTTCACTTCTGTGTCTTCCA	129							
128	TGCTTTATTGGCAACTGATATGAGCCCAAGGGGAAATTGAGAACCCCTGAAACCA	187							
130	CACCTTATTGACCTGTGATATGAGCAAGAGGAAAGTTTCAAGCGGTCCCTGAAAGCCA	189							
188	TGACATTTTCTATTATGAAAGGCCAAGAACCGGATCCAAATTAATTCAGTTTAAAGCCA	247							
190	TGACATTTCTCATTTTGAAGGCCAAGAACCGGATCCAAATTAATTCAGTTTAAAGCCA	249							
248	ATCCTCCGCTGACACTTTTGAATTAATCTGGGAGAGACAGCAACATATTTGTATGAAAC	307							
250	ACCTTCAGCTGTATCTTTCAGCTTAACGGAGAGACAGATGATTAATTTAAGATGAAA	309							
308	GGGAGGACCTTCTGTATTACCAACAGAGCCTTGGACAGGGAAACAAGATCTACTCAATC	367							
310	AGGATGACCTTCTGTATACCAACAGAGCCCTTGGACAGGAAACAAGAGCGTTCATATC	369							
368	TCCAGATTGACGCCCTGGACCTTAATGGAATTATATGTGAGGGTTCAGTCCCTATACCA	427							
370	TGCAGCTTGGACCTTGGATTTCTATGAGAGCTTAATGTGAGGGTTCAGTCCCTATACCA	429							
428	TAGAAGTGAAGGACATCAAGCAATGAGACCAAGTTTCTCAGTCAAATGACGAAGGCT	487							
430	TAGAAGTGAAGGACATCAATGAGCAACGACCAAGTTTCTCAGTCAAATGACGAAGGCT	489							
488	CAGTAAAGGCAAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTAATGCCACAGACTGG	547							
490	CAGTAAAGGCAAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTAATGCCACAGACTGG	549							
548	ATGATCCGGGACCTTCCCAATGGCCAGCTTTTATACAGATTGTATCCAGCTTCCCATGA	607							
550	ATGATCCGGGACCTTCCCAATGGCCAGCTTTTATATCAATATGTATCCCAATTTCCCAAA	609							
608	TCAACAAATGATGATCTTTTCAATGATCAACAAAGGAGGACATCTCTTATCCCGAG	667							
610	TCAACAAATGATGATCTTTTCAATGATGATCAACAAAGGAGGACATCTCTTATCCCGAG	669							
668	AGGATCTCAGGAATGAAATCTCTGATGAAGATCTCTCTTAATCTGGTGAATCTCAGTA	727							
670	AAGATCTCAGGAATGAAATCTCTGATGAAGATCTCTCTTAATCTGGTGAATCTCAGTA	729							

QY 728 AGGACATGGGAGGCGAGAGTGAATTCCTTCACTGATACCACTCTGTGATATCTAG 787
 Db 730 AGGACATGGGAGGCGAGAGTGAATTCCTTCACTGATACCACTCTGTGATATCTAG 789
 QY 788 TGACAGGAATATTTGGGAAAGACACCAAACTGTGGAGATGGTGGAAAATCTCACTGATC 847
 Db 790 TTTAGAGGAATATCTGGAAAGACACCAAACTGTGGAGATGGTGGAAAATCTGAGCTATC 849
 QY 848 CTCACCCCATCAAAATCACTCAGGTGGGATGAATGATCCGAGTGAACAATATCTTAG 907
 Db 850 CTCACCCCATCAAAATCACTCAGGTGGGATGAATGATCCGAGGAGGAGATATCTTAG 909
 QY 908 TTGACAAAGAGAGCTCCAAAGATTCCTATTTTCATTTGACCAAGAGAGATATTTACG 967
 Db 910 TCAACAGAGAGAGCTGTCCGCGTTCCATTTCTGATGACCAAGAGAGATATTTATG 969
 QY 968 TGACTCAGCCCTTGGACCGGAGAGAAAAGATGATATGTTTTTATGCAAGTTGCAAGG 1027
 Db 970 TGACTCAGCCCTTGGACCGGAGAGAAAAGATGATATGTTTTTCTTGGCAACTGCCAAGG 1029
 QY 1028 ATGAGTACGGAGAAACCACTTTCATATCGGCTGAATTCATGTAAGTTAAAGATTAA 1087
 Db 1030 ATGAGTACGGAGAAACCACTTTCATATCGGAAATTTAATGTAAGTTAATGACATTA 1089
 QY 1088 ATGATATTCACCTACATGTCCTGACACAGTAACTGATTTGAGTCCAGAGATGAAAC 1147
 Db 1090 ATGACAACTCCAGCACTGTCTGTCTCCAGTACTGTATTTGAAAGTCCAGAGATGAAAC 1149
 QY 1148 GACTGGGTAACTGATATGGGACCTTACTGCACTGACAGGATGAAGAAATATCTGCCA 1207
 Db 1150 CATTTGGGTAACTGATATGGGACCTTGTGAGGCCCATGATGAGCAAGCTTAACAATCA 1209
 QY 1208 ACAGTTTTCATTAACAGATTTGTGAGCAAACTCCCAAACTTCCCATGATGAGACTCT 1267
 Db 1210 ACAGTATTTGAAATACAGCTTTGTAGACCAACCAACCAAGTTCCCTGAGTGAATTT 1269
 QY 1268 TCTTAATTCAAACCTATGCTGGAATGTTTACAGTTAGCTTAAACAGTCTTTGAAGAGC 1327
 Db 1270 TTTCTAATGTTGAATATGAGGAGAAAGTTTCAAGTTAATTAACAGTCTTTGAAGAGC 1329
 QY 1328 ATATCTCTCACTGATCACTTAACAGATGAGTGTCTGCAAAAGATTTCAAGACCTTTGTT 1387
 Db 1330 ACAGTCTCTCACTGATCACTTAAGATGAGTGTCTGCAAGATTTCAAGACCTCTGTT 1389
 QY 1388 TTGTGAATATCAAGCTTATGATATCAATGATATGATATCCCATCTTTGAAAATTCAGAT 1447
 Db 1390 ATATTCAGTCAAGCTTATGATATCAATGATATGATATCCCATCTTTGAAAATTCAGAT 1449
 QY 1448 ATGGAACCTGACTCTTGTGAGAGACAAACAAATGGGATCCACATCTTAAACATCCAG 1507
 Db 1450 ATGGAAGCAAGACTGTGTCTGAGAGACAAACCACTGGATCCACATCTTAAATATCCAG 1509
 QY 1508 CCACTGATGCTGATGAGCCATTTACTGGGAGTTCTAAATATCTGTATCATATCATAAAG 1567
 Db 1510 CTACAGATGCTGATGAGCCCTTTAACAAGGAGCTCTAAATCTGTACAAAGATTGTACAG 1569
 QY 1568 GAGACAGTGAAGGAGCGCTGGGGTTGACACAGATCCCATACCAACCGGATATGTCA 1627
 Db 1570 GAGACAGGAGGAGGAGCTGAGAGTGTGACAGATCCCAACCAATGACGAGATATGTCA 1629
 QY 1628 TAATTAAGAAAGCTCTTGTATTTGAAAACAGCAGTGTTCACATTTGTGTCAAGAG 1687
 Db 1630 AGATCAAAAAGCTCTTGTATTTGAAAATCAACAGTGTGACAGATTTGTGTCAAGAG 1689
 QY 1688 AAAATCTGAGCCTTGAAGTTTGTGTGTAAGTAAATGCAATGCAAGTCTTTTGGCAAGTTCA 1747
 Db 1690 AAAATCTGAGCCTTGAAGTTTGTGTGTAAGTAAATGCAATGCAAGTCTTTTGGCTTGG 1749
 QY 1748 CGCTTATTTGAGAGATGTAAGTAAATGCAATTTTCCCAACAGTATTTCCAGAGCA 1807
 Db 1750 AGCTGATTTGAGAGATGTAAGTAAATGCTGTGTGTTCCTCCAGCAAGTAATTTCCAGAGCA 1809

QY 1808 AAGTCATGAGAGATGATCTATAGCACTAAAGTGGGCAATGATGATCGCCAGAGATCCAG 1867
 Db 1810 AAGTCATGAGAGATGATCTATAGCACTAAAGTGGGCAATGATGATCGCCAGAGATCCAG 1869
 QY 1868 AAGCTCTGACATTAAGCTATTTCACTGAGGGGAGACACAGAGGTTGGCTTAAATGACC 1927
 Db 1870 AAGGCTGACATTAAGCTATTTCACTGAGGGGAGACACAGAGGTTGGCTTAAATGACC 1929
 QY 1928 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1987
 Db 1930 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1989
 QY 1988 GAGTCAAGAGTGTGGGCAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2047
 Db 1990 GAGTCAAGAGTGTGGGCAAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2049
 QY 2048 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2107
 Db 2050 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2109
 QY 2108 TGTCTTCTGAGCAATCCCTCAGTGAACCTGGAAGTCTCAATTTCCAGGCTTACATGATGAT 2167
 Db 2110 TGTCTTCTGAGCAATCCCTCAGTGAACCTGGAAGTCTCAATTTCCAGGCTTACATGATGAT 2169
 QY 2168 ATCAGCACTTATTTCCGAGGATCCCATTTTACATTTTCCCTCGGAGTGAAGCTTACAA 2227
 Db 2170 ATCAGCACTTATTTCCGAGGATCCCATTTTACATTTTCCCTCGGAGTGAAGCTTACAA 2229
 QY 2228 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2287
 Db 2230 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2289
 QY 2288 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2347
 Db 2290 AAGTCACTGATGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2349
 QY 2348 TGAAGAGCACTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2407
 Db 2350 TGAAGAGCACTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2409
 QY 2408 GAGCAGAGGATCAAGACTGGGATACCACTGTGGGCAATGCAATGATGATGATGATGATGAT 2467
 Db 2410 GAGCAGAGGATCAAGACTGGGATACCACTGTGGGCAATGCAATGATGATGATGATGATGAT 2469
 QY 2468 CCACTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2527
 Db 2470 CCACTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2529
 QY 2528 AAGGCAAGATATGTTGAAGTGTCAAGCACTGAAAGTCAAACTCTGAGAGCTGAA 2587
 Db 2521 AAGGCAAGATATGTTGAAGTGTCAAGCACTGAAAGTCAAACTCTGAGAGCTGAA 2577
 QY 2588 TTTGAAAAGAAATGTTGAATTTATATGCAAGTGTATTTCA 2630
 Db 2578 TTTGAAAAGAAATGTTGAATTTATATGCAAGTGTATTTCA 2620

RESULT 14
 AF177669 3351 bp mRNA linear ROD 05-JUN-2001
 LOCUS AF177669
 DEFINITION Mus musculus liver-intestine cadherin mRNA, complete cds.
 ACCESSION AF177669
 VERSION AF177669.1 GI:5762489
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
 REFERENCES
 1 (bases 1 to 3351)
 Angres, B., Kim, L., Jung, R., Geesner, R. and Tauber, R.
 LI-cadherin gene expression during mouse intestinal development
 Dev. Dyn. 221 (2), 182-193 (2001)
 MEDLINE 21270055

PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

113/6485
2 (bases 1 to 3351)
Angres,B., Kim,L. and Tauber,R.
Direct Submission
Submitted (14-AUG-1999) Institut fuer Klinische Chemie und
Biochemie, Virchow-Klinikum der Humboldt-Universitaet zu Berlin,
Augustenburger Platz 1, Berlin 13353, Germany
Location/Qualifiers

FEATURES

source

CDS

1. .3351
/organism="Mus musculus"
/strain="Swiss Webster/NIH"
/db_xref="taxon:10090"
/tissue_type="intestine"
/dev_stage="17 day embryo"
44. .2527
/note="LI-cadherin; cell-cell adhesion molecule"
/codon_start=1
/product="liver-intestine cadherin"
/protein_id="AA051125.1"
/db_xref="GI:5762490"
/translation="MWSAQLHPLCLTLVTCGYGEKFGSPKPMSTSIIFGQEPS
QVIFQKPNPVAVFELTGETDGI FXIEKGLHYTRALDRTRAVAHLOALDSHG
AIVDGPVIRIVKIDNDNPTLOSKEYGSRONRPGKPMYNAVDLDDPATNG
QLFQVIVIOLOINDWYFOIDSKTGAISLTPGSOELDPVKYNSVLVSVDMGGO
SENSFDTTIVDISIRENTKAPBPVIRNSTDPHPIKTTVOVNDPQAOVSLVANK
KLSFPSPSIDQENIVTQALDREKSHVFFATADENKPLAFLIYKIDIND
NPTCLSPTVFEVQENEPNGSIGIFEAHDMENANNINSLIKVLDDQPKVPSGL
FLIGVEGKVLQSKKODSPQYNLSIVSDVDPEKILYKIVQDTEGRLEIVDPTN
SNYKSLSEDTAIGSTIILOATDADPEPTGSSKILYKIVQDTEGRLEIVDPTN
AGYVRIKPPDPTQPVPSIVFOANDEPPTGKIEVYASFASELEIVDPTN
QRIFOANVEDAAVGRVAVTADPEGLVYSIKKMMGWLKISVGEITSAAPL
DRSTESVYRQVAVTEVSSLSLSTADPFLHVLVDNDNPKLADLTGCFPCPLAP
GSLIFETVDDQSLRKPFTFALGREGLDSEVSKINDTHRLSTRHREEQVYN
IPRINDGQPPMEGTIVFLPVTFCQCVGSCFPAGRODGIPTVGMVAGILTPFLVI
GIIIAVVFIRMKDKVENPOSPEKPLRS"

BASE COUNT
ORIGIN

966 a 748 c 746 g 891 t

Query Match 52.3%; Score 1747.8; DB 10; Length 3351;
Best Local Similarity 80.4%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 492; Indels 13; Gaps 2;

QY 50 TTTGTGACTTCAGAAAGAAATTGAATAAGAAACATGATCTTCAGGCCCATCTTC 109
DB 4 TTTCTTACTTAAGAAAGAGTCCAGTGAAGAGC-AAGATGCTGTCTCCAGCTTC 62
QY 110 ACTCCGCTGCTCTTATGCTTATTTGGCACTGATGATGCGCAAGAGGGAAGTTTA 169
DB 63 ACTTCTGTGCTCTTCTACACTTTATTTGACCTGTGATATGGGAAAGGAAGTTCA 122
QY 170 GTGAGCCCTGAAACCCATGACATTTTCTATTATGAAGGCCAAGAACCAATTGA 229
DB 123 GCGGTCCCTGAAAGCCATGACATTTCTCATTTTGAAGGCCAAGAACCGAGTCA 182
QY 230 TATTCAGATTAAAGCCCACTCTCTGCTGATCTTTGAACCTAACTGGGAGACACA 289
DB 183 TATTCAGATTAAAGCCCACTCTCTGCTGATCTTTGAACCTAACTGGGAGACACA 242
QY 290 ACATATTGATGAACGGAGGGACTTCTGTATTACACAGAGCTTGGACAGGGA 349
DB 243 GTATATTTAAGATGAAGAGATGACCTTCTGTATCACACAGAGCCCTGACAGAGAA 302
QY 350 CAAGATCTACTCAAGATCTCCAGGTGACGCCCTGACGCTTAATGAATTATAGTAGG 409
DB 303 CAAGAGGGCTTCATCATCTGACAGCTTGACGCTTGATTTCAATGAGCTAATAGTAGG 362
QY 410 GTCCAGTCCCTATCACTAAGTGAAGACATTAAGACATGACATCCAGCTTTCTCC 469
DB 363 GTCCAGTCCCTATCACTAAGTGAAGACATTAAGACATGACATCCAGCTTTCTCC 422
QY 470 AGTAAAGTAAAGAGGCTCAGTAAAGGAGAACTCTCGCCAGGAAAGCCCTTCTTGATG 529
DB 423 AGTAAATATGAAGGCTCAGTAAAGGAGAACTCTCGCCAGGAAAGCCCTTCTTGATG 482

QY 530 TCAATGCCACAGACCTGATGATCCGGCCACTCCCAATGGCCAGCTTATACAGATG 589
DB 483 TCAATGCCACAGATCGATGATCCGGCCACTCCCAATGGCCAGCTTATACAAATG 542
QY 590 TCATCCAGCTTCCATGATCAACATGTCATGATCTTCAATGATCAACAAACGGAG 649
DB 543 TCATCCAGCTTCCCAATCAACATGTCATGATCTTCAATGATCAACAAACGGAG 602
QY 650 CCATCTCTTTACCCGAGAGGATCTCAGAAATTGAATCTCTCTAAGATCTCTCTATA 709
DB 603 CAATATCGCTTACCCAGAGAGATCCAGAAATGATCCAGTGAAGAAATCTCTCTATA 662
QY 710 ATCTGTGATCTCAGTGAAGACATGGAGGCCAGATGAGAAATCTCTCAGTATACCA 769
DB 663 ACTGTGTGCTCTGAGTGAAGACATGGAGGCCAGATGAGAAATCTCTCAGTATACCA 722
QY 770 CATCTGTGATATCATAGTGAAGAAATATTTGAAACACCAAAACCTGTGAGATGG 829
DB 723 CATATGTATATTTCTATATAGAGAAATATCTGAAACACCAAAACCTGTGAGATTA 782
QY 830 TGGAAATCTCAACTGATCTCTACCCCATCAAAATCACTCAGGTGCGGTGAGATCCGG 889
DB 783 GAGAAATCTCGACTGATCTCTACCCCATCAAAATCACTCAGGTGAGTGAATGACCCAG 842
QY 890 GTGCAATATTCCTAGTTGAAGAAAGAGCTGCAAGATTTCCATTTCAATTGACC 949
DB 843 GGGCCAGATTTCTTATGATCAACAGAGAAAGCTGTGCGCTTCCATTTCTGATGAGAC 902
QY 950 AGAAGAGATATTTACGTACTCAGCCCTTGACCCGAGAGAAAGATGATATGTTT 1009
DB 903 AGAAGAGATATTTATGATGATCAGCGCTTGACCCGAGAGAAAGATCAGATGTTT 962
QY 1010 TTTATCAGTTGCAAGATGATGATGAGAAACCACTTTCAATCTCGCTGGAATTCATG 1069
DB 963 TCTTTCACACTGCCAAGATGAGAAATGAGAAACCACTTCAATCTCAGTGAATTTATG 1022
QY 1070 TAAAGTTAAAGATTAATGATTAATCAACCTTACATGCTCGTACAGTAAACCGTATTTG 1129
DB 1023 TGAAGTTATGATTAATGATTAATCAACCAACCAATGCTGTCTCTGATGATGATTTG 1082
QY 1083 AAGTCCAGAGATGAGAACCATTTGGTAAACGATCGGGATCTTTGAGGCCCATGATATGG 1142
DB 1190 ATGAGAAATATCTGCCAAGCTTTCTAACTACAGATTTGGAGCAAACTCCCAAC 1249
QY 1143 ACGAAGCTAACCAATCAACATGATTTGAAATACAGCTTGTGACCAAAACCCAAAG 1202
DB 1250 TTCCCATGATGAGACTCTTCTTAATCAAACTATGCTGATGATGATGATGATGATG 1309
QY 1203 TTCCCTCAGTGAAGCTTTTCTCATTTGATGATGATGATGATGATGATGATGATG 1262
QY 1310 AGTCTTGAAGAGCAAGATCTCTCAGTACAACTTAAGATAGAGTGTGACAAAG 1369
DB 1263 AGTCTTGAAGAGCAAGAGCTCTCAGTACAACTTAAGATAGAGTGTGACAAAG 1322
QY 1370 ATTCAAGACCTTTGTTTGTGCAATCAAGTATGATTAATCAATGATCAATGCCCA 1429
DB 1323 ATTCAAGACCTTTGTTTGTGCAATCAAGTATGATTAATCAATGATCAATGCCCA 1382
QY 1430 TCTTTGAAAATGAGATTAAGAAACCTGCTCTGCTGAGACCAAAACATTGGGCTCA 1489
DB 1383 TCTTTGAAAATGAGATTAAGAAACCTGCTCTGCTGAGACCAAAACATTGGGCTCA 1442
QY 1490 CCATCTTAAACATCAAGGCACTGATGCTGATGAGCAATTAATGAGAGTTCATAATTC 1549
DB 1443 CCATCTTAAACATCAAGGCACTGATGCTGATGAGCAATTAATGAGAGTTCATAATTC 1502
QY 1550 TGTATCAATCATTAAGAGAGACATGAGAGAGCTCTGGGGGTTGACACAGATCCCATTA 1609
DB 1503 TGTATCAATCATTAAGAGAGACATGAGAGAGCTCTGGGGGTTGACACAGATCCCATTA 1562

Db 301 CAGCTTTTATCAAAATGTCATCCAGCTTCCAAATCAACAATGTATGTACTTCAG 360
 QY 631 ATCAACAACAAACGGAGCCATCTCTTTACCCGAGAGGATTCAGAAATGTAATCT 690
 Db 361 ATGACAAACAAACAGGGGCAATCTCCCTTACCCCAAGAGTCTCAGATATGTAATCA 420
 QY 691 GCTAAGAAATCCCTTCTTAATCTGATGATGATGATGATGATGATGATGATGATGAT 750
 Db 421 ATAAAGATCTCTTAATCAACCTGGTGTCTCTGTGAAGACATGGGGGCGCAATGAG 480
 QY 751 AATTCCTTCAGATGATCAACATCTGTGATATATCAATGATGATGATGATGATGATGAT 810
 Db 481 AATTCCTTCAGATGATCAACATCTGTGATATATCAATGATGATGATGATGATGATGAT 540
 QY 811 CCAAAACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 870
 Db 541 CCAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 871 GTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 930
 Db 601 GTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 931 TTCCCATTTTCAATTTGACCAAGAGATATTTACGTGATCTGAGCCCTTGGACCGA 990
 Db 661 TTTCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 991 GAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1050
 Db 721 GAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 1051 TATCCGCTGGAATTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1110
 Db 781 TATCCGCTGGAATTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 840
 QY 1111 TCAACAGTAAACCGTATTTGAGTCCAGAGATGATGATGATGATGATGATGATGAT 1170
 Db 841 TCTCAAGTGAATGATTTGAAGTCCAGAGATGATGATGATGATGATGATGATGATGAT 900
 QY 1171 CTTACTGACATGACAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1230
 Db 901 TTTGCTCCCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 1231 GTGGAGCAATCCCAATCTCCATGATGATGATGATGATGATGATGATGATGATGAT 1290
 Db 961 GTAGACCAACACCAAGTTCCTCAGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1291 ATGTTACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350
 Db 1021 AAGGTCAAGTTGGTAAACGGTCTTGAAGAGCAAGAGTCTCAGTACCACTTAACA 1080
 QY 1351 ATAGAGGTGTGACAAAGATTTCAAGACCTTTGTTGTGCAATCAAGTATGATGAT 1410
 Db 1081 GTAGAGGTGTGACAAAGATTTCAAGACCTTTGTTGTGCAATCAAGTATGATGAT 1140
 QY 1411 ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1470
 Db 1141 ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1471 GACACAAACATTTGGATCCACATCTTAACATCCAGGCACTGATGATGATGATGAT 1530
 Db 1201 GACACAGCATCGATCCACATCTTAATCAATCCAGGCACTGATGATGATGATGATGAT 1260
 QY 1531 ACTGGAGTTCTAAATTTCTGTATCATTAAGAGAGACAGTGAAGGAGCGCTGGGG 1590
 Db 1261 ACCGGAGCTCTAAATTTCTGTATCATTAAGAGAGACAGTGAAGGAGCGCTGGAA 1320
 QY 1591 GTTGACAGATCCCAATCCCAACACCGGATATGATGATGATGATGATGATGATGAT 1650
 Db 1321 GTTGACAGATCCCAATCCCAACACCGGATATGATGATGATGATGATGATGATGAT 1380
 QY 1651 GAAAACAGAGCTGTTCCACATGTTGTTCAAGCAGAAAAATCCTGAGCCTTATGATGTT 1710
 Db 1381 GAAAACAGAGCTGTTCAATGATGTTGTTCAAGCAGAAAAATCCTGAGCCTTATGAT 1440

QY 1711 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1770
 Db 1441 GGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1771 GAAGCACTCAATTTTCCCAACAGTATTCAGAGCAAGATGATGATGATGATGATGAT 1830
 Db 1501 GAAGTGCCTGATTTTCCCAAGCAATATTCAGAGCAAGATGATGATGATGATGATGAT 1560
 QY 1831 GGCATTAAGTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1890
 Db 1561 GGCATTAAGTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1891 CTGAGGGAGACACAGAGGTTGGTTAAATTTGACCAAGTATGATGATGATGATGATGAT 1950
 Db 1621 CTGAGGGAGACACAGAGGTTGGTTAAATTTGACCAAGTATGATGATGATGATGAT 1680
 QY 1951 GTGCTTCATTTGACAGAGAGCCGAGATCCATATCGGATACAGTGTGGC 2003
 Db 1681 ACTGCTCCGCTGACAGAGGAAACAGAAAGTGTGTATGATGATGATGATGATGAT 1733

Search completed: June 21, 2003, 23:43:29
 Job time : 8339 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:07:55 ; Search time 4860 Seconds
(without alignments)
4982.209 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIIQAHLSICLIMLYLATG.....DKGKNVESQAASEVKEPLRS 832

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DBY=xlp
-Q=/cgm21/USFTO.spool/US10025380/runac_17062003_173328_24668/app_query.fasta_1.1031
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US10025380 @CGN 1.1 3745 @runac_17062003_173328_24668 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_hcg_hum.*
31: em_hcg_in.*
32: em_hcg_other.*
33: em_hcg_mus.*
34: em_hcg_pln.*
35: em_hcg_rtd.*
36: em_hcg_man.*
37: em_hcg_vrt.*
38: em_sy.*
39: em_hcg_hum.*
40: em_hcg_mus.*
41: em_hcg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4321	100.0	2499	6 140307	140307 Sequence 2
2	4321	100.0	2499	6 181197	181197 Sequence 2
3	4321	100.0	3345	6 AX193509	AX193509 Sequence
4	4321	100.0	3345	6 AX330272	AX330272 Sequence
5	4321	100.0	3345	6 AX332252	AX332252 Sequence
6	4321	100.0	3345	6 AX410642	AX410642 Sequence
7	4321	100.0	3345	9 HSU07969	U07969 Human Intes
8	4303	99.6	3654	6 AX330312	AX330312 Sequence
9	4303	99.6	3654	6 AX335495	AX335495 Sequence
10	4303	99.6	3654	6 AX411163	AX411163 Sequence
11	4303	99.6	3654	6 HSRNALIC	X83228 H. sapiens m
12	3450	79.8	3351	10 AF177669	AF177669 Mus muscu
13	3450	79.8	3414	10 D87912	D87912 Mus musculu
14	3414	79.0	3628	10 RNCADHER	X78997 R. norvegicu
15	2328	53.9	1733	5 RAPPDRPT	L46874 Rattus norv
16	1427.5	33.0	3373	5 AF428098	AF428098 Dantio rer
17	954	22.1	2846	4 OCUT8945	U28945 Oryctolagus
18	922.5	21.3	2764	9 AF016272	AF016272 Homo sapi
19	922.5	21.3	2820	9 BC027912	BC027912 Homo sapi
20	908.5	21.0	2848	6 AX092366	AX092366 Sequence
21	908.5	21.0	2848	6 AX358934	AX358934 Sequence
22	908.5	21.0	2848	6 AX362427	AX362427 Sequence
23	878.5	20.3	2864	10 BC015251	BC015251 Mus muscu
24	878.5	20.3	2927	10 AF016271	AF016271 Mus muscu
25	878.5	20.3	2938	6 AX338110	AX338110 Sequence
26	877.5	20.3	2493	6 AX338112	AX338112 Sequence
27	862.5	20.0	2799	9 AK093905	AK093905 Homo sapi
28	758.5	17.6	2516	6 AX396194	AX396194 Sequence
29	715	16.5	2954	5 CHKRCG	DI4459 Chicken mRN
30	709.5	16.4	2827	5 AF418565	AF418565 Dantio rer
31	705	16.3	3230	10 MMRCCADA	X69966 M. musculus
32	705	16.3	3274	10 MUSRC	DI4888 Mouse mRNA
33	699.5	16.2	2853	9 HSNCAD	X54315 Human mRNA
34	699.5	16.2	4132	9 S42303	S42303 N-cadherin
35	699	16.2	3690	5 S82457	S82457 Xenopus mat
36	696.5	16.1	3203	5 GGNCAD	X07277 Chicken mRN
37	694.5	16.1	3269	10 BC022107	BC022107 Mus muscu
38	694.5	16.1	4321	6 AX305854	AX305854 Sequence
39	694.5	16.1	4321	10 AB008811	AB008811 Mus muscu
40	694.5	16.1	4321	10 MUSCADNA	M31131 Mouse neutra
41	691.5	16.0	3448	9 HUMNCAOH	M34064 Human N-cad
42	688.5	15.9	3839	4 BTCDHN	X53615 B. taurus mR
43	687.5	15.9	3310	9 AK091496	AK091496 Homo sapi
44	686.5	15.9	3036	10 AF097593	AF097593 Rattus no
45	686.5	15.9	4350	10 AB017695	AB017695 Rattus no

RESULT 1

ALIGNMENTS

140307
 LOCUS 140307 2499 bp DNA linear PAT 13-MAY-1997
 DEFINITION Sequence 2 from patent US 5620855.
 ACCSSION 140307
 VERSION 140307.1 GI:2082599
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE Unclassified.
 1 (bases 1 to 2499)
 AUTHORS Danczig, A.H., Hoskins, J.A. and Skatrud, P.L.
 TITLE Mammalian influx peptide transporter
 JOURNAL Patent: US 5620855-A 2 15-Apr-1997;
 FEATURES
 source Location/Qualifiers
 1..2499 /organism="unknown"
 BASE COUNT 728 a 559 c 570 g 642 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2499
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-10-025-380-1081 (1-832) x 140307 (1-2499)
 QY 1 MetIleuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTYrLeuAlaThrGly 20
 DB 1 ATGTAATCTTCAGGCCCATCTTCACTCCCTGTCCTCTTATGCTTTATTTGGCAACTGGA 60
 QY 21 TyrGlyGlnGlyGlyPheSerGlyProLeuLysProMetThrPheSerIleTYrGlu 40
 DB 61 TATGGCAAGAGGGAGGAGTTAGTGACCCCTGAACCCATGACATTTTCTATTATGAA 120
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAspProAlaValThrPhe 60
 DB 121 GGCCAGAACCGAGTCAATATATATCCAGTTTAAGCCAAATCCCTGCTGTACTTTT 180
 QY 61 GluLeuThrGlyGlyLysThrAspAsnIlePheValIleGluArgGlyGlyLeuTYrTYr 80
 DB 181 GAATTAATGGGGAGACAGACACATATTTGTGATAGAACGGAGGACTTCTGTATTAC 240
 QY 81 AsnArgAlaLeuAspArgGlyLysThrArgSerThrHisAsnLeuGlnValAlaIleLeuAsp 100
 DB 241 AACAGAGCTTGACAGAGGAAACAAGATCTACTACAAATCTCCAGTTGCAGCCCTGGAC 300
 QY 101 AlaAsnGlyIleIleValGlyGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB 301 GCTATATGAATTAATAGTGAGGGGTCCATCTCTATCCACATAGAAAGTAAAGACATCAAC 360
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTYrGlyGlySerValArgGlnAsnSerArg 140
 DB 361 GACAATGACCCCAAGCTTTCTCCAGTCAAGTACAGAGCTCAGTAAAGGACAACTCTCCG 420
 QY 141 ProGlyLysProPheLeuTYrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 DB 421 CCAGAGAAAGCCCTTCTGTATGTCAATGCCACAGACTGATGATCCGGCCACTCCCAAT 480
 QY 161 GlyGlnLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTYrPhe 180
 DB 481 GGCCAGCTTTATATACCAATGTCATCCAGCTTCCCATGATCAACAATGATATGACTTT 540
 QY 181 GlnIleAsnAsnLysThrGlyValAlaIleSerLeuThrArgGlyGlySerGlnGlnLeuAsn 200
 DB 541 CAGATCAACAACAAACGGAGCCATCTCTCTTACCCAGAGGGATCTCAGAAATTGAAT 600
 QY 201 ProAlaLysAsnProSerTYrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
 DB 601 CTGTGTAAGAAATCCTCTATATATCTGTGATCTCAGTGAAGAGACATGGAGGCCAGAGT 660

QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTYrLys 240
 DB 661 GAGAAATTCCTTCAGTATACCAATCTGTGATATATAGTACAGAGATATTGGA 720
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 DB 721 GCACCAAAACCTGTGAGATGGGAAACCTCAACGATCTCCACCCATCAAAACACT 780
 QY 261 GlnValArgThrAsnAspProGlyAlaGlnTYrSerLeuValAspLysGlyLysLeuPro 280
 DB 781 CAGGTCCGGGAGATATATCCGGTGCAATATCTTCTTGTTGACAAAGAGAGCTGCCA 840
 QY 281 ArgPheProPheSerIleAspGlnGlyAspIleTYrValThrGlnProLeuAspArg 300
 DB 841 AGATTCCTCATTTTCAATTGACCGAGAGAGATATTAGTGACTCAGCCCTTGGACCGA 900
 QY 301 GlyGlyLysAspAlaTYrValPheTYrAlaValAlaLysAspGlyTYrGlyLysProLeu 320
 DB 901 GAAGAAAGATGACATATGTTTTTTATGCAAGTTGCCAAGAGATGATCGAAGAACACACT 960
 QY 321 SerTYrProLeuGlnIleHisValLysValLysAspIleAsnAspAspProThrCys 340
 DB 961 TCATATCCGTGGAAATTCATGTAAAGTTAAAGATTTAATGATATCCACTTACATGT 1020
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB 1021 CCTCACCAAGTACCTTATTTAGAGTCCAGAGAAATGAAGACATGGGTAAACAGTATCGGG 1080
 QY 361 ThrLeuThrAlaHisAspArgAspGlyGluAsnThrAlaAsnSerPheLeuAsnTYrArg 380
 DB 1081 ACCCTTACGACATACACAGAGGATGAAGAAATACGCCAACAGTTTCTTAACTACAG 1140
 QY 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTYrAla 400
 DB 1141 ATGTGAGCAAACTCCCAACTTCCATGAGAGGACTCTTCTTAATCCAACTATGCT 1200
 QY 401 GlyMetLeuGlnLeuValLysGlnSerLeuLysLysGlnAspThrProGlnTYrAsnLeu 420
 DB 1201 GGAATGTACAGTTACTTAAACAGTCTTGAAGAACAGATATCTCTCGTCAACTTA 1260
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 DB 1261 ACGATAGAGGTGTCACAAAGATTCAAGACCCCTTGTGTTGTGCAAAATCAACGTTATT 1320
 QY 441 AspIleAsnAspGlnIleProIlePheGlyLysSerAspTYrGlyAsnLeuThrLeuAla 460
 DB 1321 GATATCAATGATATGATCCCATCTTGAAGAAATCAGATTATGGAATCTGACTTGGCT 1380
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1381 GAAGACAAACAATTTGGGTCCACATCTTAACCATCAGGCCCATGATGCTATGAGCCA 1440
 QY 481 PheThrGlySerSerLysIleLeuTYrHisIleIleLysGlyAspSerGlyAlaGlu 500
 DB 1441 TTTACTGGAGATTCTAAATATCTGTATCATATATAAAGGAGACAGTGAAGGACCCCTG 1500
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTYrValIleIleLysLysProLeuAsp 520
 DB 1501 GGCGGTGACACAGATCCCATCCATCCACACCGGATATGTCAATTAATAAGCCCTTGAT 1560
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 DB 1561 TTGGAACAGACAGCTTTTCCACATTTGTCTTCAAGGAGAGAAATCTGAGCCTTATG 1620
 QY 541 PheGlyValLysTYrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB 1621 TTGGGTGAGAGTACATGCAATGCTTTTGGCAAGTTCCAGCTTATGACAGATGTG 1680
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB 1681 AATGAAGCACCTTAATTTTCCCAACAGTATTCAGGAAATCATGAGTGGAGATGTAGCT 1740
 QY 581 IleGlyThrLysValGlyAsnValThrAlaAspProGlyLysLeuAspIleSerTYr 600

[illegible]

Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatch:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-025-380-1081 (1-832) x I81197 (1-2499)			
QY	1 MetIleLeuGlnAlaHisLeuHisSerLeuCySLeuLeuMetLeuTyrLeuAlaThrGly	20	
Db	1 ATGATACTTCGAGGCCCATCTTCACCTCCGTCGTCTTTATGCTTTATTTGGCAACTGGA	60	
QY	21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerTLeTyrGlu	40	
Db	61 TATGGCCAAAGAGGGGAAAGTTTGTGTGACCCCGAAACCACATGACATTTTCTATTATGAA	120	
QY	41 GlyGlnGluProSerGlnTLeIlePheGlnPheLysAlaAsnProAlaValThrPhe	60	
Db	121 GGCCAAAGACCGAGTAAATTATATTCCAGTTTAAAGCCCATCTCTCGCTGTACITTT	180	
QY	61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyr	80	
Db	181 GAACATACTGGGAGACAGACAACATTTGTGATGAAAGCGAGGAGCTTCGTATTAAC	240	
QY	81 AsnArgAlaLeuAspArgLysThrArgSerThrHisAsnLeuGlnValAlaLeuAsp	100	
Db	241 AACGAGCCTTGGACAGGGAAACAGATCTACTCAATCTCCAGTTGCGACCCCTGGAC	300	
QY	101 AlaAsnGlyLeIleIleValGluGlyProValProIleThrIleGluValAspIleAsn	120	
Db	301 GCTAATGAAATTATAGTGAAGGGGTCCAGTCCCTATACCATGAAATGAAGACATCAAC	360	
QY	121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg	140	
Db	361 GACATATGACCCGACGTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAAGCAACATCTGC	420	
QY	141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspAspProAlaThrProAsn	160	
Db	421 CCAAGAAAGCCCTTCTGTATGATCAATGCCAAGACCTGAGTATCCCGGCACTCCCAAT	480	
QY	161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetCileAsnAsnValMetTyrPhe	180	
Db	481 GGCCAGCTTATTAACAAGATGTGCATCCAGCTCCCATGATCAACAAATGATATGACTT	540	
QY	181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn	200	
Db	541 CAGATCAACAACAAGCGGAGCCATCTCTTACCCAGAGGAGATCTCAGGAATTGAT	600	
QY	201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer	220	
Db	601 CTGCTAAGATATCTTCTATATATCGTGATCTCAGTGAAGAGACATGGAGGCGCAAGT	660	
QY	221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleThrLys	240	
Db	661 GAGAAATTCCTTCGTATACCAATCTGTGGATCATCAATGACAGAAATATTGGAAA	720	
QY	241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr	260	
Db	721 GCACCAAAACCTGTGAGATGGTGGAAAACTCAATGATCTCCACCCCATTAATAATCACT	780	
QY	261 GlnValaLysGTrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro	280	
Db	781 CAGGTGGGCGAATGATCCCGGTCAACAATTCCTAGTTGCAAAAGAAAGAACTCCCA	840	
QY	281 ArgPheProPheSerTLeAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg	300	
Db	841 AGATTCCTCATTTCAATTACCAAGAGAGATATTTCAGTACCTCAAGCCCTTGGACGA	900	
QY	301 GlnGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu	320	
Db	901 GAAGAAAAGATCAATATGTTTTTTTATTCAGAGTGCMAAGATGATGCGAAAACCACTT	960	
QY	321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProProThrCys	340	

Db	961	TCATATCCGCTGGAAATTCATGTAAATCTTAAAGTATTAATGATATATCCACCTACAGT	1020
Qy	341	ProSerProValThrValPheGluValGlnGluAsnGluValGluGlyAsnSerIleGly	360
Db	1021	CCGTCACACAGTAACCGGTATTGGAGTCCAGGAGAAATGAACACACTGGGTAAACAGTATCCGG	1080
Qy	361	ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnThrArg	380
Db	1081	ACCTTACTGTGACATGACAGGAGATGAAGAAATATCTGCACACAGTTTCTTAACTCACGG	1140
Qy	381	IleValGluGlnThrProLeuLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla	400
Db	1141	ATTGGAGGACAACTCCCAACTTCCATGATGACCTTCTCTTAATCCAAACGTATGCT	1200
Qy	401	GlyMetLeuGlnLeuAlaGlyGlnSerLeuTyrGlyAsnAspThrProGlnTyrAsnLeu	420
Db	1201	GGAAATGTACAGTGTACTAAACAGCTCTTGAAGAAAGCAAGATACTCCTCAGTACAACTTA	1260
Qy	421	ThrIleGluValSerAspIleAspPheLeuPheLeuTyrPheValGlnIleAsnValIle	440
Db	1261	ACGATAGAGGGTCTGCACAAAGATTTCAGAACCCCTTTGTGTGGCAATCAACCTTAT	1320
Qy	441	AspIleAsnAspGlnIleProIlePheGluTyrSerAspTyrGlyAsnLeuThrLeuAla	460
Db	1321	GATATCAATGATCAGATCCCATCTTTGAAAATACGATTAATGAGAACTCACTTCTCT	1380
Qy	461	GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro	480
Db	1381	GAAAGCACAACATCGGGTCCACCATCTTAACCATCCAGCCAGCCAGTGTGATGAGCCA	1440
Qy	481	PheThrGlySerSerLeuIleLeuTyrHisIleIleLeuGlyAspSerGluGlyArgLeu	500
Db	1441	TTTACTGGAGATTCTTAAATCTCTGATCATATACATAAAGGAGACAGTAGGAGAGCTGTG	1500
Qy	501	GlyValAspThrAspProHisIleThrAsnThrGlyTyrValIleIleLeuTyrProLeuAsp	520
Db	1501	GGGGTTGACACAGATCCCATTCACACACCGGATATGTCATATTAATAAGCCTTGTAT	1560
Qy	521	PheGluThrAlaAlaValSerAsnIleValPheTyrAlaGluAsnProGluProLeuVal	540
Db	1561	TTTGAACACAGACGCTGTTCCAACTGTGTTCAAGACAGAAATCTGAGCTCTTAGTG	1620
Qy	541	PheGlyValIleTyrAsnAlaSerSerPheAlaIlePheThrLeuIleValIleThrAspVal	560
Db	1621	TTTGGTGTGAAGTACAAATGCAGAGTCTTTTGGCAAGTTCACGCTTATGTGACAGATGTG	1680
Qy	561	AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaTyrValSerGluAspValAla	580
Db	1681	AATGAAAGCCTCCAAATTTTCCCAACAGTATTCAGACGAAAGTACAGTAGGATTAAGT	1740
Qy	581	IleGlyThrIleValGlyAsnValThrAlaTyrAspProGluGlyLeuAspIleSerTyr	600
Db	1741	ATAGCACAATAAGTGGGCAATGTGACTGCAAGATCCAGAAAGTCTGGACATAAGTAT	1800
Qy	601	SerLeuAlaArgIleAspThrArgGlyTyrPheLeuTyrIleAspHisValIleGlyGluIlePhe	620
Db	1801	TCACTGAGGGGAGACAAAGAGTGTGCTTAATAATTGACCAAGTACTGTGTAGATCTTT	1860
Qy	621	SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGluValAlaIleThr	640
Db	1861	AGTGTGGCTCAATTGACAGAAACCCGGAACTCCATTCGGGTACAAAGTGTGGCCACA	1920
Qy	641	GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal	660
Db	1921	GAAAGTAGGGGGGTCTTCTCTTAAGCTCTGTGTACAGATTCCACCTGATCTTATGGAATGTG	1980
Qy	661	AsnAspAsnProProArgLeuAlaTyrAspTyrThrGlyLeuPhePheCysHisProLeu	680
Db	1981	AATGACAACCTCCACAGGCTACCAAGACTACAGGGGCTTGTCTTGTGCATCCCTTC	2040
Qy	681	SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGly	700

Db	2041	AGTCACCTGGAAAGTCTCATTTTGGAGGCTACTGATGATGATGACACTTATTTGGGGGT	2100
Qy	701	Prothiaphenrphreserleuglysercglyserleuglnhnaprtpgluvalserlys	720
Db	2101	CCCCATTATTAATTTTCCCTCGGAGTGGAGAACTTACAAACGACTGGAGATTTCNAA	2160
Qy	721	lleaanglythrhisalaarglserthrarghisrtharaphegluvalargalaty	740
Db	2161	ATCATAGTACTCATGCGGACTCTCTACACGACACAGACTTTTGAGGAGAGGGCGTAT	2220
Qy	741	valvalleuilearglleaenapgllylaryppropoleuugllyllevalserleu	760
Db	2221	GTCCCTCTTGATCCCAATCAATGATGGGGGTTCGGCCACCTTGGAAGGACTGTTCYTTA	2280
Qy	761	Provalrhrphecysarcyvalgluglysercyssphearpgproalaglyhisglthr	780
Db	2281	CCAGTATCAATTCCTCAGTGTGTGTGGAGAAAGTTGTTCCGGCCAGCGGTACACAGACT	2340
Qy	781	glylleprothrvalgllymetalaivalglylleleuethrthrleuvalillegly	800
Db	2341	GGGATACCCACTGTGGGCGATGGCGAGTTGTATCTGCGACCACTTCGTGTATTTGCT	2400
Qy	801	llelleleualaivalpheeliarygllyelvalysaplysglylyasapnaivalglu	820
Db	2401	ATAATTTTAGCAGTTGGTTTATCCGCAATAAAGAGATTAAGCAAAAGATTAATGTTGAA	2460
Qy	821	SeralaglnalasergluvallypProleuargser	832
Db	2461	AGTGCTCAAGCATCTGAAGTCMAACCTCTGAGAAAGC	2496
RESULT 3			
LOCUS	AX193509	3345 bp	DNA
DEFINITION	Sequence 1076 from Patent WO0149716.	linear	PAT 15-AUG-2001
ACCESSION	AX193509		
VERSION	AX193509.1	GI:15211449	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Enkaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3345) Xu,J., Lodes,M.J., Secrist,H., Benson,D.R., Meagher,M.J., Stolk,J.A., King,G.E., Wang,T. and Jiang,Y., Compounds for immunotherapy and diagnosis of colon cancer and methods for their use Patent: WO 0149716-A 1076 12-JUN-2001; CORIXA CORPORATION (US) location/Qualifiers 1..3345 /organism="Homo sapiens" /db_xref="taxon:9606"		
FEATURES	source		
BASE COUNT	951 a	746 c	720 g
ORIGIN		928 t	
Alignment Scores:			
Pred. No.:	0	Length:	3345
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-025-380-1081 (1-832) x AX193509 (1-3345)			
Qy	1	MettlleuuglnahisaleuHiserleucysleuLeuMettleyrleuAlathrgly	20
Db	88	ATGAACTTCAGGGCCCATCTTCACTCCCTGTCCTCTTAAGCTTATTTGGCACTGGA	147
Qy	21	TyrglynglngllylyspheSerclYProleuylsPrometthrPheSerleryglu	40
Db	148	TATGGCCAAAGGGGAAGTTTAGTGGAGCCCTGAAACCAAGACATTTTCTATTATGAA	207

QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe 60
 Db 208 GGCCAAAGACCGATCAATTAATATTCAGTTTAAGCCCAATCTCTGCTGTACTTTT 267
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80
 Db 268 GAACATACTGGGGAGACAGACAAACATATTTGTGATTAACGGGAGGAGCACTTGTATTAAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrIleAsnLeuGlnValAlaAlaLeuAsp 100
 Db 328 AACGAGACCTTGGACAGGAAACAAAGATCTACATCTCCAGCTTGACGCTTGAC 387
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db 388 GCTAATGAGATTATTAAGTAGAGGGTCCAGTCCCTATCAACCATAGAAAGGAGCATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg 140
 Db 448 GACATTCGACCCAGCTTCTCCAGTCAAGTACGAGGCTCAGTAAAGGCAAGACTCTGCG 507
 QY 141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspProAlaThrProAsn 160
 Db 508 CCAAGAAAGCCCTTCTGTATGTCAATGCCACAGACTGGATGATCCGGCCACTCCCAAT 567
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGCACGCTTATTAACGATTTGTATCAGCTTCCTCATGATCAACATGATGATTAATT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 628 CAATCAACAAACAAAGGAGGAGCATCTCTTACCCGAGAGGATCTCAGGAATGAT 687
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 Db 688 CCGCTAAGATCTCTTCTATATCTGATCTCAGTGAAGGACATGGAGGCGACAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
 Db 748 GAGAAATCTTCAAGTATCAACATCTGTGATATCATAGTACAGAGAAATTTGGAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 808 GACACAAACCTGTGGAGATGGGAAACCTCAATCATCTCAACCCCAACAAATCACT 867
 QY 261 GlnValArgThrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 868 CAGGTGGGTGATGATCCCGGTGACAAATATCTTAACTTGAACAAAGAGAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
 Db 928 AGATTCCTCATTTCAATTTGACCAAGAGAGATATTTAGCTGACTGAGCCCTTGACCGA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 Db 988 GAAGAAAGAGATCATATGTTTATTCAGTTGCAAGGATAGTACGGAACCAACCTT 1047
 QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
 Db 1048 TCAATATCCCTGGAATTCATGTAAAGATTAATATATATATCACTACATATGT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1108 CCGTCAACAGTAACCGATTTGAGGTCCAGAGAAATGAAGATCGGTAAACATATCGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1168 ACCCTTAACGACATGACAGGAGTGAAGAAATATCTGCCAACAGTTTCTTAACCTACAG 1227
 QY 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAAACTTCCATGATGAGACTCTTCTTAATCCAAACCTATGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420

Db 1288 GGAATGTACAGTTAGTAAACAGCTCTTGAAGACAGATATCTCTCAGTACACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1348 ACATATGAGGTGTCTCAACAAAGATTTCAAGACCTTGTGTTGTGAAATCAACGTTAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATAGATCCCATCTTGAATAATCAATATATGAAACCTGACTTCTGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAGACAAACAACTGTGGTCCACCATTTAAACATCAAGGCCACTGATGTGAGGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLeuGlyLysAspArgGluArgLeu 500
 Db 1528 TTTACTGGGAGTCTAAATTTCTGTATCATATCATAAAGGAGACAGTGAAGGACCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATACCAACCGGATATGTCTAATTAATAAAGCTCTTGAT 1647
 QY 521 PheGluThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAACACGACGCTGTTCCAACTGTGTCAAGGAAAGAAATCTGAGCCTTAGTG 1707
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGTGTGAGATCAATCAAGTCTTTTGGCAAGTTCAACGCTTATTTGACAGATGATG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1768 AATGAAGCACTTAATTTTCCAAACGATTTCAAGCAAGATCAAGTGAAGTGAAGT 1827
 QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluLysLeuAspIleSerTyr 600
 Db 1828 ATAGGCACTTAAGTGGGCAATGATGATCGCAAGAGATCCGAAGGTCTGGACATTAAC 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyLysIlePhe 620
 Db 1888 TCACCTAGGGGAGACCAAGAGGTTGGCTTAAATGACACAGCTGACTGTGAGACTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1948 AGTGTGCTCCATTGACACAGAGCCGGAAGTCCATATCGGTACAGAGTGGGCCACA 2007
 QY 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAGTAGGGGGGCTCTTAAAGCTGTGTGAGAGTTCCACTGATCTTATGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCCGAGCTACCAAGAGCTACACGCGCTTGTCTTGCCATCCCTTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2128 AGTGACCTGGAAGTCTCAATTTTCAGGCTACGATGATGATCAGACTTAATTTGCGGG 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2188 CCCCATTTTACATTTTCCCTCCGAGTGAAGCTTACAAAGCATCGGAGAACTTTCCAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db 2248 ATCAATGATACATCAAGCCGACGTCTACAGGCAACAGACTTTGAGGAGAGGGGTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyLysIleValSerLeu 760
 Db 2308 GTGCTTGTATCCGATCATATATGAGGGGTGCGCAACCTTGAAGGCAATTTGCTTTTA 2367
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780

Db 2368 CCACTTACATTCGCACTTGTGTGAGAGAAAGTTGTTCCGGCCAGCAGGACCACT 2427
Qy 781 GYIYLeProThraValGlyMetAlaValGlyIleLeuLeuThrThreLeuValIleGly 800
Db 2428 GGGATACCCCACTGCGGCGATGCGAGTGTGTACTGTGACCACTTCTGTGTGATGCT 2487
Qy 801 ILeIleLeuAlaValIlePheIleArgIleIleGlySerPheGlyGlySerPheValIle 820
Db 2488 ATATATTTTACAGTGTGTGTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAA 2547
Qy 821 SerAlaGlnAlaSerGlyValIleValProLeuArgSer 832
Db 2548 AGTCTCAAGCATCTGAAGTCAACCTCTGAGAAAC 2583

RESULT 4
AX330272 3345 bp DNA linear PAT 09-JAN-2002
LOCUS Sequence 781 from Patent WO0194629.
DEFINITION AX330272
ACCESSION AX330272
VERSION AX330272.1 GI:18103250
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horvigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 781 13-DEC-2001;
FEATURES
source 1. 3345
Location/Qualifiers
BASE COUNT 951 a 746 c 720 g 928 t
ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3345
Score: 4321.00 Matches: 832
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-025-380-1081 (1-832) x AX330272 (1-3345)

Qy 1 MetIleLeuGlnAlaHisIleLeuHisSerLeuCySteLeuMetLeuTyTrLeuAlaIleThrGly 20
Db 88 ATGATTAATTCAGGCCCACTTCTGCTCTTCTTATGCTTTATTTGGCACTGGA 147
Qy 21 TyGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyGly 40
Db 148 TATGGCCAGAGGGGAGAGTTTATGTCAGCCCTGAAACCATGACATTTTCATTTATGAA 207
Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValIleThrPhe 60
Db 208 GGGCAAGAACCGATCAATATTAATTCAGTTTAAAGCCATCTCTCGCTGACCTTT 267
Qy 61 GlnLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyTr 80
Db 268 GAATTAATCTGGGAGACAGACATATTTGTATGATGAAGGAGGAGCTTCTGTATAC 327
Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaIleAsp 100
Db 328 AACGAGCTTGGACAGGAGAAACAAGATCTACATCTCCAGTTGCAAGCCCTGGAC 387
Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIleAspIleAsn 120
Db 388 GCTATGGAATTTATATGAGAGGCTCACTCCCTATCACCATGAAGTGAAGACATCAAC 447

Qy 121 AspAsnArgProThrPheLeuGlnSerLysTyGlyGlySerValArgGlnAsnSerArg 140
Db 448 GACATCGAACCCAGCTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAAGCAACTCTCG 507
Qy 141 ProGlyLysProPheLeuTyTrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db 508 CCAGGAAAGCCCTTCTTGTATGTCATGCAATGCGACAGACTGATGATCGGCACTCCCAAT 567
Qy 161 GlyGlnLeuTyTrTyGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyTrPhe 180
Db 568 GGGCACTTATTAATCCAGATGTCATCCAGCTTCCATATCAACATGATGATCTT 627
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGluLeuAsn 200
Db 628 CAGATCAACAAACAAACGGAGGCCATCTCTTACCCGAGAGGAGATCTCAGGAATGAAAT 687
Qy 201 ProAlaLysAsnProSerTyTrAsnLeuValIleSerValIleAspMetGlyGlyGlnSer 220
Db 688 CTGTATGAGATCTCTTCTTATATCTGTGATCTCATGATGATGATGATGATGATGATGAT 747
Qy 221 GlnAsnSerPheSerAspThrThrSerValAspIleIleValIleThrGlnAsnIleTyPlys 240
Db 748 GAGAAATCTCTCAATATCCAGATCTGATGATATCATATGATGATGATGATGATGATGAT 807
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProIlePheIleThr 260
Db 808 GCACCAAACTGTGAGATGTCGAAACCTCAACATGATCTCAACCCATCAAAATCACT 867
Qy 261 GlnValArgTrpAsnAspProGlyAlaGlnTySerLeuValAspLysGlyLysLeuPro 280
Db 868 CAGGTGCGGAGAAATGATCCCGGTGCAATATCTCTTATGATGATGATGATGATGATGATGAT 927
Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyTrValIleThrGlnProLeuAspArg 300
Db 928 AGATTCATTTCAATATGACGAGAGGATATTTACCTGATGATGATGATGATGATGATGATGAT 987
Qy 301 GlnGlyLysAspAlaTyTrValIlePheTyTrAlaValAlaIleAspGlyLysTyGlyLysProLeu 320
Db 988 GAGGAAAGATGATCAATATGTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1047
Qy 321 SerTyProLeuGlnIleHisValIleValIleAspIleAsnAspAsnProProThrCys 340
Db 1048 TCATATCCGCTGGAATTCATGTAAGATTAAGATTAATGATATATCACTCACTACATGT 1107
Qy 341 ProSerProValIleThrValPheGluValGlnGluAsnGluArgGlyLysSerIleGly 360
Db 1108 CCGTCACAGTAAACCGATTTAGTTCAGGAGAGATGATGATGATGATGATGATGATGATGAT 1167
Qy 361 ThrLeuThrAlaHisAspArgGluGluAsnThrAlaAsnSerPheLeuAsnTyArg 380
Db 1168 ACCCTTATCTGACATGACAGGATGAAGAAATACGCGCAAGCTTTCTTAACTACAGG 1227
Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyTrAla 400
Db 1228 ATTGTGAGCAAACTCCCAAACTCCCATGATGATGATGATGATGATGATGATGATGATGAT 1287
Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyTrAsnLeu 420
Db 1288 GGAATGTTACAGTTAGTAAACAGCTCTTGAAGAGCAAAATACCTCTGATCACTTAA 1347
Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCySerPheValGlnIleAsnValIle 440
Db 1348 ACCATAGAGTGTCTGACAAAGATTCACAAACCTTTGTTTGTGCAAAATCAAGCTTAT 1407
Qy 441 AspIleAsnAspGlnIlePheIlePheGlnLysSerAspTyTrGlyAsnLeuThrLeuAla 460
Db 1408 GATATCAATGATCAGATCCCATCTTTGAAATAATCAATATGAAACCTGACCTTGTCT 1467
Qy 461 GlnAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db 1468 GAAGACCAAAACATTTGGGTCCACCATCTTAAACATCCAGGCGCATGATGCTGATGAGCCA 1527
Qy 481 PheThrGlySerSerLysIleLeuTyHisIleIleLysGlyAspSerGlnGlyArgLeu 500

```

Db      1528  TTTAAGGAGGAGTTCTTAAATTCGTATCAATCAATTAAGGAGACAGTGGAGGCGCTG 1587
QY      501    G1ValAspThrAspProHisThrAsnThyGlyTyrValIleIleIleValysProLeuAsp 520
Db      1588  GGGGTTCACACAGATCCCATCCATCCACACCGGATATGTCATTAATTAAGAGCCCTTGAT 1647
QY      521    PheGluThrAlaAlaValSerAsnIleValPheValSAlaGlnunProGluProLeuVal 540
Db      1648  TTTGAAAACAGAGCTGTTTCCATCATGTTGTTAAACAGAAAATCTGAGCCCTAGT 1707
QY      541    PheGlyValIleYsTyrAsnAlaSerSerPheAlaIlePheThrIleuIleValIThrAspVal 560
Db      1708  TTTGGTGTGAAGTACAAATGCAAGTCTTTTGCCAAAGTTCACGCTTATTTGACAGATGTG 1767
QY      561    AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIleValSergIleuAspValAla 580
Db      1768  AATGAAACCTCAATTTTCCACACAGATTCACACGAAAGTCACTGAGAGATGATGCT 1827
QY      581    IleGlyThrIleValGlyAsnValThrAlaIleAspProGluGluIleuAspIleSeryTyr 600
Db      1828  ATAGGCACTAAAGTGGGCAATGTGATCCAGAGATCCAGAGAGTCTGACATAGCTAT 1887
QY      601    SerIleuArgGlyAspThrArgIleTyrPLeuIleAspHisValThrGlyIleuIlePhe 620
Db      1888  TCACTGAGGAGAGACACAGAGGTTGGCTTAAATGACACGATGAGTGGATCTTT 1947
QY      621    SerValAlaProIleuAspArgGluAlaGlySerProTyrArgValGlnValValAlaThr 640
Db      1948  AGGTGTGCTCATTTGACAGAGAAAGCCGAAAGTCCATATGCGGATCAAGGTGGCCACA 2007
QY      641    GluValGlyIleYsSerSerIleuSerySerValSergIleuPheIleuIleuMetAspVal 660
Db      2008  GAAGTAGGGGGGCTTCTTAAGCTGTGTGACAGATTCACCTGATCTTAAGAGATGTG 2067
QY      661    AsnAspAsnProProAlaGleuAlaIleAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db      2068  AAAGAACAAACCTCCAGAGCTAGCCAGAGACTACACGGGCTTGTCTTCCATCCCTCC 2127
QY      681    SerAlaProGlySerIleuIlePheGluAlaThrAspAspAspGlnHisleuPheArgGly 700
Db      2128  AGTGCACCTGGAGTCTCATTTTTCGAGGCTATGATGATGATGACATTAATTTCCGGGT 2187
QY      701    ProHisPheThrPheSerIleuGlySeryGlySerIleuGlnAsnAspTyrGluValSerys 720
Db      2188  CCCCATTTTACATTTTCCCTCGGCAAGTGAAGAGCTTACAAACGACTGGGAAGTTCCAA 2247
QY      721    IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluIleuArgAlaTyr 740
Db      2248  ATCAATGTAATCATGATGCGGAGCTGTCTACAGGACACACAGACTTTGAGGAGAGGCGTAT 2307
QY      741    ValValLeuIleLeuArgIleAsnAspGlyIleArgProProLeuGluGluIleValSeryLeu 760
Db      2308  GTGCTCTGTATCCGATCAATGATGAGGAGTGGCCACCTTGGAGAGGACTTTGTTCTTTA 2367
QY      761    ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db      2368  CCAAGTACATTTCTGAGCTGTGTGGAAGGAAGTTGTTTCCGCGCAGAGAGTTCACCAACT 2427
QY      781    GlyIleProThrValGlyMetAlaValGlyIleIleuIleuThrThrLeuLeuValIleGly 800
Db      2428  GGGATATCCCACTGTGGGATGAGGAGTGTGATCTGTGACACACCTTCTGGGATTTGGT 2487
QY      801    IleIleIleuAlaValAlaPheIleArgIleIleValIleValAspIleGlyIleAspAsnValGln 820
Db      2488  ATATATTTAGCAATGTGTATATCCGATTAACAAAGAGATTAAGCAAAAGATATGTTGAA 2547
QY      821    SerAlaGlnAlaSerGluValIleYsProLeuArgSer 832
Db      2548  AGTGCTCAAGCATCTGAAGTCAAACTCTGAGAAAGC 2583

```

```

LOCUS      AX332252          3345 bp      DNA      linear      PAT 09-JAN-2002
DEFINITION Sequence 2761 from Patent WO0194629.
ACCESSION  AX332252
VERSION     AX332252.1  GI:18122886
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
             Horrigan, S., Soppet, D. R. and Weaver, Z.
TITLE       Cancer gene determination and therapeutic screening using signature
             gene sets
JOURNAL     Patent: WO 0194629-A 2761 13-DEC-2001;
             Avalon Pharmaceuticals (US)
FEATURES
             Location/Qualifiers
             source          1..3345
                             /organism="Homo sapiens"
                             /db_xref="taxon:9606"
BASE COUNT  951 a 746 c 720 g 928 t
ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3345
Score:          4321.00  Matches:      832
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    100.00%  Indels: 0
DB:              Gaps: 0

US-10-025-380-1081 (1-832) x AX332252 (1-3345)
QY      1    MetIleleuGlnAlaHisIleuHisSerIleuCySleuLeuMetLeuTyrIleuAlaThrGly 20
Db      88    ATGATACCTTCAGGCCCATCTTCATCCCTCGTGTCTTTATGCTTAATTTGGCACTGGA 147
QY      21    TYGylGlnGluGlyIleYsPheSerGlyProLeuIleProMetThrPheSerIleTyrGlu 40
Db      148    TATGGCCAAAGAGGAGAAAGTTAGTGAACCCCTGAACCAATGACATTTCTATTATGAA 207
QY      41    GlyGlnIleProSergIleIlePheGlnPheIleValIleGluArgGluGlyLeuIleTyr 60
Db      208    GGCCAAGAACCGAGTCAATTTATTTCCAGTTTAAGGCCAATCCCTGCTGTGACTTTT 267
QY      61    GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuIleTyr 80
Db      268    GAACATACTGGGGAGACAGACAAATATTTGTGATGAACGGGAGGAGCTTCTGTATTAC 327
QY      81    AsnArgAlaIleuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaIleuAsp 100
Db      328    AACAGAGCTTGGACGGGAAACAGATCTACTACAAATCTCAGGTTGACGCTCGAGAC 387
QY      101    AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIleAspIleAsn 120
Db      388    GCTAATGGAATTAATGATGAGGAGTCCAGTCCCTTAATCAATGAAGTGAAGACATCAAC 447
QY      121    AspAsnArgProThrPheLeuGlnSeryIleTyrGluGlySerValArgGlnAsnSeryArg 140
Db      448    GACAAATCGACCAAGTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAAGGACGAATCTCCG 507
QY      141    ProGlyIleProPheLeuTyrValAlaAsnAlaThrAspIleuAspAspProAlaThrProAsn 160
Db      508    CCAAGAAACCTCTTCTGTATGATGACATGACAGACCTGAGATGATCGGCACTCCCAAT 567
QY      161    GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db      568    GGCACGCTTATTAACAAGTTGTATCATCCAGCTTCCATGATCAACATGTCTATGTAATT 627
QY      181    GlnIleAsnAsnIleThrGlyAlaIleSeryLeuThrArgGluGlySeryGlnIleuAsn 200
Db      628    CAGATCAACAAACAAACGGGAGCCATCTCTTAACCCGAGAGGAGATCTCAGAAATTAAT 687

```

QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 DB 688 CCGCTTAAGAACTCTTCTATTAATCTGATCTCAAGTGAAGCACTGGAGGCCAAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrPys 240
 DB 748 GAGAAATTCCTTCACTGATACCAATCTGTGGATATCACTAGTACAGAGAAATATTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleValIleThr 260
 DB 808 GCACCAAAACCTGTGGAGATGTGGAAACCTCACTATCTCTCAACCCCACTCAAAATACT 867
 QY 261 GluValArgTrpAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 DB 868 CAGGTGGGTGGATGATCCCGGTGCACATATTCCTTAGTGAACAAAGAACTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluLysAspIleTyrValThrGlnProLeuAspArg 300
 DB 928 AGATTCACATTTTCAATTGACACAGAAAGAGATATTACGTGACTCAGCCCTTGACCGA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 DB 988 GAAAGAAAGATGATGATGTTTATGCAAGTTCAGATGAGTACGGAAACCACTT 1047
 QY 321 SerTyrProLeuGluIleHisValLysAspIleAsnAspAsnProProThrCys 340
 DB 1048 TCATATCCGCTGGAAATTCATGTAAGTTAAAGATTAATGATTAATCCACTACATGCT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB 1108 CCGTCAACAGTAACGATATTGAGGTCCAGAGAAATGACGATCGGGTAACAGATATCGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 DB 1168 ACCCTTACTGCACATGACAGGATGAAGAAATATCTGCCAACATTTTCTTAACCTACAGG 1227
 QY 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 DB 1228 ATTGTGAGACAAATCTCCAAATCTCCATGATGATGACTCTTCTTAATCCAAACCTATGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
 DB 1288 GGAATGTTACAGTAAACAGTCTCTGAAAGACAAAGATATCTCCCTCACTACACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 DB 1348 ACGATTAAGGTGTCTGACAAAGATTTCAGACCCCTTGTGTGCAAAATCAAGTTATT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 DB 1408 GATATCAATGATCGATCCCATCTTGTAAATACTAGATTATGAAACCTGACTCTTGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAAGACAAACATTTGGGTCCACCATCTTAACATCCAGGCCACTGATGATGAGCA 1527
 QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluLysArgLeu 500
 DB 1528 TTTTACTGGAGATTCTAAATCTGTATCATATCAATAAGGAGACAGTGAAGGACGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 DB 1588 GGGGTGTGACAGATCCCATTAACCAACCCGGAATGTCAATTAATTAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 DB 1648 TTTTAAACAGACGCTGTTTCCACATGTTGTTTCAAGAGAAATTCCTGAGCCTTAGTG 1707
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB 1708 TTTGTGTGAAGTACATGCAAGTCTTTTCCCAAGTTCACGCTTATTTGTACAGATGTG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580

DB 1768 AATGAAGCACTCAATTTTCCCAACAGCTATTCAGACGAAAGTCAAGTAGATAGCT 1827
 QY 581 IleGlyThrLysValAlaLysAsnValThrAlaLysAspProGluLysLeuAspIleSerTyr 600
 DB 1828 ATAGCACTAAAGTGGGCAATGATGATGCTCCAGACTCCAGAAAGTGTGACATTAACCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyIleIlePhe 620
 DB 1888 TCACCTAGGGGAGACACAAAGAGTTGCTTAATTAATGACCACTGACTGGTGAATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 DB 1948 AGTGTGCTTCATTTGACAGAGAAAGCCGGAAGTCCATATCGGGTACAAAGTGGCCACA 2007
 QY 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 DB 2008 GAAGTAGGGGGGCTTCTTAAGCTCTGTGTGCAAGTTCACCTGATCTTATGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPheCysHisProLeu 680
 DB 2068 AATGACACACCTCTCCAGGCTAGCCAGACTACACGGCTTGTTCTTCCCTCCCTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
 DB 2128 AGTGCACCTGGAAGTCTCATTTTCCAGGGTACTGATGATCAGCACTTATTCGGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluValSerLys 720
 DB 2188 CCCCATTATTACATTTTCCCTCGGACGTGAAAGCTTCAAAACACCTGGGAAGTTTCCAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 DB 2248 ATCAATGTGACTATGCTCCGACACTGTCTTACCAGCACACAGACTTTAGAGAGAGCGGTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 DB 2308 GTGCTGTGATCCGATCATATGATGGGGTCCGCCACCTTGGAAAGCATTTGTTCTTTA 2367
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 DB 2368 CCAATTACATCTGCAAGTGTGTGGAAGAGATGTTTCCGGCAGCAGGTCCACAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 DB 2428 GGAATACCACTGTGGGATGAGTGTGTTATCCGATTAAGAGATTAAGCAAAAGTATGTTGAA 2547
 QY 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 DB 2488 ATATTTTACAGTGTGTTATCCGATTAAGAGATTAAGCAAAAGTATGTTGAA 2547
 QY 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 DB 2548 AGTGTCCAGCATCTGAAGTCAAACTCTGAAGAGC 2583
 DB 2548 AGTGTCCAGCATCTGAAGTCAAACTCTGAAGAGC 2583

RESULT 6
 AX410642 3345 bp DNA linear PAT 14-JUN-2002
 LOCUS Sequence 3289 from Patent WO0229103.
 DEFINITION AX410642
 ACCESSION AX410642
 VERSION AX410642.1 GI:21443347
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 TITLE Gene expression profiles in liver cancer
 JOURNAL Patent: WO 0229103-A 3289 11-APR-2002;
 FEATURES
 location/Qualifiers
 source 1..3345

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="EMBL/GenBank Accession No. U07969"

BASE COUNT 951 a 746 c 720 g 928 t

ORIGIN

Alignment Scores:

Pred. No.:	0	Length:	3345
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-025-380-1081 (1-832) x AX410642 (1-3345)

QY 1 MetIleuGlnAlaHisIleuHisSerLeuCYsIleuLeuMetLeuTYrLeuAlaThrGly 20
 DB 88 ATGATACCTTCAGGCCCACTTCCTGCTCCCTGCTCTTATGCTTTATTTGGCAACTGGA 147
 QY 21 TYrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTYrGlu 40
 DB 148 TATGGCCAAAGGGGAAGTTTAGTGAGACCCCTGAAACCATGACATTTCTATTATGAA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
 DB 208 GGCCAAAGACCGAATCAATTAATATTCAGTTTAAAGCCAAATCCTCGCTGTCATTTT 267
 QY 61 GlnLeuThrGlyGluThrAspAsnIlePheValIleGluLysGluGlyLeuLeuTYrTYr 80
 DB 268 GAATTAACCTGGGAGACAGACAAACATATTTGTATTAACGGAGGAGGACTTGTATATAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB 328 AACAGACCTTTGGACAGGAAACAAGATCTACTACATCTCCAGTTTGCAGCCCTGAC 387
 QY 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB 388 GCTAATGGAATTAATAGGAGGGGTCAGTCCCTATACCAATGAAAGGAAAGGACATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTYrGlyGlySerValArgGlnAsnSerArg 140
 DB 448 GACAAATGACCCAGGTTCTCCAGTCAAGTACGAAAGCTCAGTAAAGGCAACTCTCCG 507
 QY 141 ProGlyLysProPheLeuTYrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 DB 508 CCAAGAAAGCCCTTTCTTGTATGTAATGCAAGACCTGGATGATCCGGCCACTCCCAAT 567
 QY 161 GlyGlnLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTYrPhe 180
 DB 568 GGCACAGCTTATTACCAAGATGTCATCAGCTCCAGTCAATGATCAATGTCATGTACTTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 DB 628 CAAATCAACAACAAGAGGAGCCATCTCTTAACCCGAGAGGAGATCTCAGGAATGAAT 687
 QY 201 ProAlaLysAsnProSerTYrAsnLeuValIleSerValLysAspMetGlyGlnGlnSer 220
 DB 688 CCGCTAGAAATCTTCTTAATCTGTATGATCTCAATGAAAGCAATGGAGGCAAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTYrLys 240
 DB 748 GAGAAATCTTCACTGATACCAATCTGTGGAATATCAATGACAGGAATATTGGA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 DB 808 GCACCAAAACCTGTGGAGATGGTGAACCTCAATCATCTCAACCCATCAAAATCACT 867
 QY 261 GluValArgThrAspAspProGlyAlaGlnTYrSerLeuValAspLysGluLysLeuPro 280
 DB 868 CAGGTGGCTGGAAATGATCCCGGTGCAATATTTCTTAGTGAACAAGAGAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTYrValThrGlnProLeuAspArg 300

DB 928 AGATTCACATTTCAATTCAGCAGAAAGAGATTTATCGTACTGACCTTCGACCGCA 987
 QY 301 GluGluLysAspAlaTYrValPheTYrAlaValAlaLysAspGluTYrGlyLysProLeu 320
 DB 988 GAAAGAAAGGATCATATGTTTTTTATGACGTTGCAAGGATAGTACGAAACCACTT 1047
 QY 321 SerTYrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
 DB 1048 TCATTCGGCTGGAAATTCATGTAAGTTAAAGTAAATTTATATATATCACTACATGT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlnLysSerIleGly 360
 DB 1108 CCGTCAACAGTAAACCGATTTGAGTCCAGAGAAATGAACGATCGGTTACGATTCGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgGluGluAsnThrAlaAsnSerPheLeuAsnTYrArg 380
 DB 1168 ACCCTTAATGACATATGACAGGATGAAAGAAATACGCCAAAGTTTCTTAACTACAG 1227
 QY 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTYrAla 400
 DB 1228 ATTGTGAGCAAACTCCCAACTTCCCATGATGACTCTTCTTAATCCAAACTATGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTYrAsnLeu 420
 DB 1288 GGAATGTTACAGTTAGCTAAACAGTCTTGAAAGAACAGATACTCCTCAGTACAACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCYsPheValGlnIleAsnValIle 440
 DB 1348 ACATATGAGGTGTCTCAACAAGATTCAGAACCTTTGTTGTGCAAACTCAAGTTAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTYrGlyAsnLeuThrLeuAla 460
 DB 1408 GATATCAATGATCAAGATCCCATCTTTGAAATTCAGATTAATGAAACCTGACTTTGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAGACAAACAATGGGTCCACCATCTTAACCATCAGGCCACTGATGTATGAGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuTYrHisIleIleLysGlyAspSerGluGlyLeu 500
 DB 1528 TTTACTGGGAGTTCTAAATTTCTGATCATATCAATTAAGGGAACAGTGGAGCAGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTYrValIleIleLysLysProLeuAsp 520
 DB 1588 GGGGTGACACACATCCCATCAACACCGGATATGTCATTAATTAAGGCTCTTGAT 1647
 QY 521 PheGluThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 DB 1648 TTTGAAACAGCAGCTGTTTCCAAATTTGTTCAAGCAAGAAATCTGAGCTCTAGTG 1707
 QY 541 PheGlyValLysTYrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB 1708 TTTGGTGTAAAGTACATCAAGTTCTTTTGGCCAAAGTCAACGCTTATTTGACAGATG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB 1768 AATGAAGACCTCAATTTTCCAAACGATTTCAAGCAAGAACTCAGTGAAGTATGCT 1827
 QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTYr 600
 DB 1828 ATAGGCACTTAAGTGGCAATGATGACTGCCAAGATCCGAAGGTCTGGACATTAACAT 1887
 QY 601 SerLeuArgGlyAspThrArgIleTYrPheLysIleAspHisValThrGlyLysIlePhe 620
 DB 1888 TCACTAGGGGAGACCAAGAGGTGGCTTAAATTAATGACACAGTGACTGGTGAATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluLysSerProTYrArgValGlnValAlaThr 640
 DB 1948 AGTGTGCTTCATTTGACAGAGAGCCGGAAGTCCATATCGGGTAAAGGTGGGCGACA 2007
 QY 641 GluValGlyLysSerSerLeuSerValSerGluPheHisLeuIleLeuMetAspVal 660

Db 2008 GAAGTAGGGGGCTCTTCTTAAGCTCTGTGTCAGAGTTCCACGTCATCTTATGAGTGTG 2067
 Qy 661 AaaaapaaanProProaHgluAlaLysaPtyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATATCAACCTCCAGAGCTTAGCCAGAGACTACAGGGCTGTGTTCTTCTGCACACCCCTC 2127
 Qy 681 Ser1aProGlySerLeuIlePheGluAlaThrAspAspGlnIleLeuPheArgGly 700
 Db 2128 AGTCACCTGGAACTTCATTTTCAGAGCTTCTATGATGATGACACTTATTTCCGGGCT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2188 CCCCATTTTTCATTTTCCCTCGGAGTGGAGAGTTTACAAACGACTGGAGAGTTTCCAAA 2247
 Qy 721 IleAsnGlyThrHisAlaIleArgLeuSerThrArgHisThrAspPheGluGluValArg 740
 Db 2248 ATCAATGATCTCATGCGCCGACTGTCTACACGACACAGACTTTGAGAGAGGGCCGTAT 2307
 Qy 741 ValValLeuIleAlaGlyIleAsnAspGlyGlyValArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTGCTCTTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAGAGGACTTGTCTTTA 2367
 Qy 761 ProValThrPheCysSerCysCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATCTTCAGATTGTGTGTGAGAGAGTTGTTCCGGCAGCAGGTCAACAGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCCACTGGGGGATGAGCAGTTGTACTGCTGACACACCTTCTGTGATTTGT 2487
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysIleAspLysGlyLysAspAsnValGly 820
 Db 2488 ATAAATTTTTCAGATGTGTGTGTTATTCGCTAATAAGATTAAGCAAGATTAATTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2548 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAAAC 2583
 RESULT 7
 HSU07969 3345 bp mRNA linear PRI 15-JUN-1994
 LOCUS Human intestinal peptide-associated transporter HPT-1 mRNA,
 DEFINITION complete cds.
 ACCESSION U07969
 VERSION U07969.1 GI:483391
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 3345)
 AUTHORS Dantzig,A.H., Hobkins,J., Tabas,L.B., Bright,S., Shepard,R.L.,
 Jenkins,I.L., Duckworth,D.C., Sportsman,R., Mackensen,D.,
 Roestek,P.R. and Skelrud,P.L.
 TITLE Association of intestinal peptide transport with a protein related
 to the cadherin superfamily
 JOURNAL Science 264 (5157), 430-433 (1994)
 MEDLINE 94204643
 PUBMED 8153632
 2 (bases 1 to 3345)
 Roestek,P.R.
 REFERENCE Direct Submission
 AUTHORS Submitted (24-MAR-1994) Paul R. Roestek Jr., Lilly Research
 JOURNAL Laboratories, Eli Lilly and Co., Lilly Corporate Center,
 Indianapolis, IN 46285, USA
 FEATURES
 source 1..3345
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="Caco-2"
 /cell_type="adenocarcinoma"
 /tissue_type="colon"
 88..2586

/codon_start=1
 /product="intestinal peptide-associated transporter HPT-1"
 /protein_id="AAA19021.1"
 /db_xref="GI:483392"
 /translation="MTLQAHLSLCLMLNLYATGVCQGEKFSGLPKMTFSYEGEP
 SQTLEQNPAPAVTFELTGETDNIIVIEREGLLYYNRALDRSTRSHNLQVALDAN
 GLIVEGPAVITTEVKDINDNRPFLOSKEGSRQMSRGEPLVYNAIDLPAPN
 GOLXYQIVLQPMINVMYFQINNKTKASLFRGSGQSELPANPSYNIIVSYKMG
 OSNBSRSDTSVDIIVTENIWKAPKEVENVNSTDPPIKIQVRNNDGQASLVKD
 EKLPREPSIDDEGDIYVQPLDREKDAVFAVAKDERGKLSYELIHVQKDN
 DNEPTSPSPVTFEVEENRLENSIGTLTKADRENTANSELTATVIECTPLPMDG
 LEPLQTYAGMLQAKSLKKODTPQNLITEVSDKQFKTLCEVQINVIDINDIPLFE
 KSDYGNLTLEADTNISITLTQATDADPEFGSSKILYHIKGSDEGTLGVDPDT
 NNGVYIIEKPLDETAAVSNIVFAKNEPEPLVEGVYKNSSEFAKFLIYDVVEAEP
 NGSVPAKXSEVDAIGTKVNTAKDPEGLDISYLRGDRGMLKIDHTGEIFSVAP
 LDREASPRRVQVVAITEVGSSLSYSVERHLLIMVNDNPPRLAKDYGLFCHPLSA
 PSLILEATYDDDLFRGHFTFSLSGSLQNDMEVSKNGTRARLSTRHTPEEAY
 VVLRINDGRPEPLGVSLPVTFCVSGSCRPGRGHQGTLPVGMAGVILTLVLV
 IGIILAVFIRIKDKGKDNVSAQSEVYKPLRS"
 BASE COUNT 951 a 746 c 720 g 928 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-10-025-380-1081 (1-832) x HSU07969 (1-3345)
 Qy 1 MetIleleuGlnAlaHisleuHisserLeuCysleuLeuMetLeuTyrleuAlaThrGly 20
 Db 88 ATATATCTTCAAGCCCACTTCACTCCCTGTGCTTCTTATGATTTTGGCAACTGGA 147
 Qy 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGly 40
 Db 148 TATGGCCAGAGGGGGAAGTTTACTGGACCCCTGAACCCATGACATTTCTATTATGAA 207
 Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
 Db 208 GGCACAGAACCGAGTCAATATATTCAGTTTAAGCCCATCTCTGCTGAGACTTTT 267
 Qy 61 GluLeuThrGlyGlyLysPheAsnIlePheGlnPheLysAlaAsnProProAlaValThrPhe 80
 Db 268 GAACCTAACCTGGAGAGACACATATTTGTGATGAAAGGAGGAGGACTTCTGTATTAC 327
 Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 328 AACAGAGCCTTGGACAGGGAACAGATCTACTCACATCTCCAGGTTGACGCCCTGGAC 387
 Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db 388 GCTAATGAAATTATAGGAGGGGTCCAGTCCCTATACCCATTAAGAGGAGGACATCAAC 447
 Qy 121 AspAsnArgProThrPheLeuGlnSerLeuTyrGlyLysSerValArgGlnAsnSerArg 140
 Db 448 GACATATGACCAACGTTTCTCCAGTCAAAATAGAGAGGCTCAAGTAAAGGCAACTCTCGC 507
 Qy 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 508 CCAAGAAAGCCTTCTTGTATGTCAATGCCACAGACCTGATATATCGGCCACATCCCAAT 567
 Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGCACGCTTATTAACGATGTGATCCACCTTCCCATGATCAACAATGTCAGTACTCTTT 627
 Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 628 CAGATTAACAACAACAAAGGAGGCACTCTTCAACCGAAGGAGATTCAGAAATTAAT 687
 Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnGlnSer 220

Db 688 CCTCTAGAGATCCCTTCTATTAATCTGTGATCTCAGTAGAGGACATGGGGGCGCAGAGT 747
 Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTrpLys 240
 Db 748 GAGAAATCTTCAATGATACCAATCTGTGATATCTAGTACAGAGAAATATTGGAAA 807
 Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 808 GCACCAAAACCTGTGAGATGGTGGAAAACCTCACTGATCTCTCAACCCCATCAAAATCACT 867
 Qy 261 GluValArgTrpAsnAspProGluValAlaGlnTrpSerLeuValAspLysGluLysLeuPro 280
 Db 868 CAGGTGGGTGGAGATGATCCCGGTGCACAATATTCCTTAGTTGACAAAGAGAACCTCCCA 927
 Qy 281 ArgPheProPheSerIleAspGlnGluValAspIleTrpValThrGlnProLeuAspArg 300
 Db 928 AGATTCCCATTTTCAATTGACACGAGAGAGAGATATTTACGTGACTCAGCCCTTGGACCGA 987
 Qy 301 GluGluLysAspAlaTrpValPheTrpAlaValAlaLysAspGluTrpGlyLysProLeu 320
 Db 988 GAAAGAAAAGATGATATGTTTATTCAGTTGCAAGATGAGTACGGGAAAACCACTT 1047
 Qy 321 SerTrpProLeuGluIleHisValLysValLysAspIleAsnAspAsnProTrpThrCys 340
 Db 1048 TCATATCCGTGGAAATTCATGTAAAGTTAAAGATTATGATTAATCCACCTACATGAT 1107
 Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGluLysAsnSerIleGly 360
 Db 1108 CCGTACACAGTAAACCGATTGAGTGCAGAGATACACATCGGGTAAACAGATCGGG 1167
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTrpArg 380
 Db 1168 ACCCTTACTGACATGACAGGAGTGAAGAAAATCTGCCAAGCTTTTCTTAATCAACG 1227
 Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTrpAla 400
 Db 1228 ATTGTGAGCAAACTCCCAAACTCCCATGATGATGATCTCTCAATCCCAAACTGATGCT 1287
 Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTrpAsnLeu 420
 Db 1288 GGAATGTTCACTTACCTTAACACGTCTTGAAGAGAAATCTCTCTGATCAACTTA 1347
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1348 ACCATAGAGGTGCTGACAAAGATTTCAAGACCTTTGTTTGCAAAATCAAGTTATT 1407
 Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTrpGlyAsnLeuThrLeuAla 460
 Db 1408 GATATCAATGATCAATCCCATCTTTGAAATAATCAGATTATGGAACCTGACTCTTGGT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAAGACAAACATGGGTCCACCATCTTAACATCCAGGCCACTGATGTGATGAGCCA 1527
 Qy 481 PheThrGlySerSerLysIleLeuTrpHisIleIleLysGlyAspSerGluGlyArgLeu 500
 Db 1528 TTTACTGGGAGTTCTAAATTTCTGTATCATATCAATAAGGAGACAGTGGAGGCCCTG 1587
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTrpValIleIleLysLysProLeuAsp 520
 Db 1588 GGGGTTCGACACATATCCCATACCAACACCGGATATGCTAATTAATAAAGCCCTTGAT 1647
 Qy 521 PheGluThrAlaIleAlaSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAAACGACACTGTTTCCACATTTGTGTTCAAGCAAAAATCCTGAGCCTTAGTG 1707
 Qy 541 PheGlyValLysTrpAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGGTGTGAAGTACATGCAATGTTCTTTGGCAAGTTCAGCTTATTTGACAGATGTG 1767
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580

Db 1768 AATGAAGACCTCAATTTTCCCAACGTAATTCACGAAAGTCACTAGTAGTACT 1827
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTrp 600
 Db 1828 ATAGGACCTAAAGTGGCAATGTGACTGCCAAGATCCAGAAAGTCTGGACATAGTACT 1887
 Qy 601 SerLeuArgGlyAspThrArgGlyTrpLeuLysIleAspHisValThrGlyLysIlePhe 620
 Db 1888 TCACGTAGGGAGACACAAAGAGTGGCTTAATAATGACCAGCTGACTGGTAGACTTTT 1947
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTrpArgValGlnValAlaThr 640
 Db 1948 AGTGTGGCTTCATTGACAGAGAGCCGGAAGTCCATATCGGTACAAAGTGGGCCACA 2007
 Qy 641 GluValGlyLysSerSerLeuSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAATAGGGGGGCTTCTTCACTTCACTGTGTCAAGATTCACCTGATCTTATGAGATGTG 2067
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTrpThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCCAAGGCTAGCCAAAGACTACACGGGCTGTCTTCCATCCCTTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2128 AGTGCACCTGGAAAGTCTCAATTTTCAGAGCTACTGATGATGATCAGCATTAATTCGGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGluAsnAspTrpGluValSerLys 720
 Db 2188 CCCCATTTTACATTTTCCCTCGGACGTGAAGCTTCAAAAACATCGGAAGATTCCAAA 2247
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTrp 740
 Db 2248 ATCAATGTACTCATATCCCGACGTCTTACCAAGCACAACAGACTTTAAGAGAGGGCTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTCTCTTGTATCCGATCATATGATGGGGGTCCGCCACCTTGGAAAGCAATGTTCTTTA 2367
 Qy 761 ProValThrPheCysSerCysValGlyLysSerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTAATCTTGCAGTGTGTGGAAGAGATGTTTCCGCGACAGATCCACGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
 Db 2428 GGGATACCACTAGTGGGCAATGGCAGTGGTATATCTGACACACCTTCTGGATTTGGT 2487
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2488 ATTAATTTTACAGTTGTGTTTATCCGATTAAGAAAGATTAAGGCAAAAGATATGTGA 2547
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuAspSer 832
 Db 2548 AGTGTCAAGCATCTGAAGTCAAACTCTGAAGAAC 2583
 RESULT 8
 AX330312 3654 bp DNA linear PAT 09-JUN-2002
 LOCUS Sequence 821 from Patent WO0194629.
 DEFINITION AX330312
 VERSION AX330312
 KEYWORDS
 SOURCE
 ORGANISM
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
 Horrigan, S., Soppet, D.R. and Weaver, Z.
 TITLE
 Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL
 Patent: WO 0194629-A 821 13-DEC-2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 Location/Qualifiers

source 1..3654
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1052 a 804 c 777 g 1021 t
ORIGIN

Alignment Scores:
Prod. No.: 1,99e-317 Length: 3654
Score: 4303.00 Matches: 828
Percent Similarity: 99.76% Conservative: 2
Best Local Similarity: 99.52% Mismatches: 2
Query Match: 99.58% Indels: 0
DB: Gaps: 6

US-10-025-380-1081 (1-832) x AX330312 (1-3654)

QY 1 MetlleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
DB ATGATCTTCAGGCCCATCTTCACTCCCTGTGCTCTTATGCTTTATTGGCACTGGA 152

QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
DB TATGGCCAGAGGGGAAATTAGTGAACCCCTGAACCCATGACATTTCTATTATGAA 212

QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAspProAlaValThrPhe 60
DB GGCCAGAACCCGAGTCAATATTATTCAGTTTAAAGCCCATCTCTGCTGACTTTT 272

QY 61 GluLeuThrGlyGlnThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyr 80
DB 273 GAATTAATGGGGAGACAGACAAATATTTGTGATAGAACGGAGGAGACTTGTATTAC 332

QY 81 AsnArgAlaLeuAspArgGlnThrArgSerThrHisAsnLeuGlnValAlaIleAsp 100
DB 333 AACGAGCCCTTGAACAGGAAACAAAGATCTACATCTCCAGGTGACGCCCTGAC 392

QY 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGlnValLysAspIleAsn 120
DB 393 GCTATGAAATTATAGTGAAGGCTCCAGTCCCTTACCATTAAGTGAAGACATCAAC 452

QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlnGlySerValArgIleAsnSerArg 140
DB 453 GACATGACCCCACTTCTCCAGTCAAAAGTACAAAGGCTCAGTAAAGCACTGCGC 512

QY 141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspProAlaThrProAsn 160
DB 513 CCAGAAAGCCCTTCTGTATGATCAATGCCAGACCTGATGATCCGGCCACTCCAAAT 572

QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
DB 573 GGCCAGCTTATTAACAGATTGTATCCAGCTTCCAGTATCAACATGTCTATCTTT 632

QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlnGlySerGlnIleAsn 200
DB 633 CAGATCAACAACAAACAGGAGCCATCTCTTACCCGAGAGGATCTCAGAAATTGAAT 692

QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
DB 693 CTGCTAGAAATCTTCTCTATATCTGTGATCTCAGTGAAGACATGGAGGCGCAAGT 752

QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
DB 753 GAGAAATCTTCAATGATACCAATCTGTGATATCATAGTGAAGAAATTTTGGAAA 812

QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
DB 813 GCACCAAAACCTGTGAGATGTGGAAACCTCACTGATCTCAACCCCATCAAAATCACT 872

QY 261 GlnValArgTyrPheAspProGlyValAlaGlnTyrSerLeuValAspLysGlnLysLeuPro 280
DB 873 CAGGTGGGTGGAATGATCCGGGTGCACAATATTTCTTATGTTGACAAAGAACTGCA 932

QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300

DB 933 AGATCCCAATTTCAATTGACCCAGAAAGATATTAGTGACTGAGCCCTTGGACCG 992
QY 301 GlnGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGlnTyrGlyLysProLeu 320
DB 993 GAAGAAAGATGACATATGTTTTTATATGATGATGCAAGAGATGATGCAAGAAACACTT 1052

QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
DB 1053 TCATATCCCTGGAATTCATGTAAGAATTAAAGATTAATATGATATATCACTACATGAT 1112

QY 341 ProSerProValThrValPheGlnValGlnGluAsnGlnArgLeuGlyAsnSerIleGly 360
DB 1113 CCCTACACAGTAACCGTATTGTAGTCCAGAGAAATGAACGACTGGTAACTATACGAG 1172

QY 361 ThrLeuThrAlaHisAspArgAspGlnGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
DB 1173 ACCCTTAACGACATACAGAGGATGAAGAAATACGCAACGTTTCTTAACTACAGG 1232

QY 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
DB 1233 ATTGTGAGCAACTCCCAACTTCCCATGATGACTCTTCTTATCCAAACTATGCT 1292

QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysLysGlnAspThrProGlnTyrAsnLeu 420
DB 1293 GGAATGTTACAGTTACTTAAACAGTCTTGAAGAACCAAGATACTCTCAGTCAACTTA 1352

QY 421 ThrIleGlnValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
DB 1353 ACATATAGAGTGTCTCAACAAAGATTCAGAACCCCTTGTGTGCAAACTCAAGTATAT 1412

QY 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAla 460
DB 1413 GATATCAATGATCAGACCCCATCTTGAATAATCAGATTATGAAACCTGACTTTGCT 1472

QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
DB 1473 GAAGACAAACATGTGGTCCACCATCTTAAACATCCAGGACCTGATGATGAGCCA 1532

QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGlnGlyArgLeu 500
DB 1533 TTTACTGGAGTCTTAAATTTGTATCATATATGAAGGAACAGTGGAGGACGCTG 1592

QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
DB 1593 GGGGTGACACATATCCCATACCAACCGGATATGTCTATTAATAAGCCTCTTGAT 1652

QY 521 PheGlnThrAlaAlaLysSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
DB 1653 TTTGAAACAGCAGCTGTTTCAACATTTGTCTCAAGCAGAAATCCTGAGCCTTAGTG 1712

QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
DB 1713 TTTGGTGTGAAGTACATGCAAGTCTTTTGGCAAGTCCAGCTTATTTGGACAGATGTG 1772

QY 561 AsnGlnAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
DB 1773 AATGAGCACTCAATTTTCCCAACAGTATTTCAAGCGAAATGCTGAGATGATGCT 1832

QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGlnGlyLeuAspIleSerTyr 600
DB 1833 ATAGGCACTAAAGTGGCAATGTGACTGCAAGATTCGAAGAGTGTGACATTAAGCTAT 1892

QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
DB 1893 TCACTGGGGGAGAACCAAGAGTGGCTTAAATGACACATGATGTGTGAGATCTTT 1952

QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
DB 1953 AGTGTGCTCATTTGACAGAGAACCGGAAGTCCATATGTGGGTAAAGTGTGGCCACA 2012

QY 641 GluValGlyLysSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspVal 660

Db 1173 ACCCTTACTGACATGACAGGATGAGAAATATCTCCACAGCTTTCTTAACTACAGG 1232
 Qy 381 TLevalGluGlnThrProLysLeuPheMetAspGlyLeuPheLeuIleGlnThrTyra 400
 Db 1233 ATTGTGAGGAAATCCCAACCTTCCATGATGATGATCTTCTTAACTCAACCTATGCT 1292
 Qy 401 GlyMetLeuGlnLeuAlaValGlnSerLeuLysLeuGlnPheThrProGlnTyraPheLeu 420
 Db 1293 GGAATGTTCAGTAAAGTAAACAGTCTTGAAGAGAGAAATCTCTCTGATCAACTTA 1352
 Qy 421 ThrIleGlnValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 Db 1353 ACCGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGCAATCAAGCTTATT 1412
 Qy 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrglyAsnLeuThrLeuAla 460
 Db 1413 GATATCAATGATGACACCCCATCTTGAATAATCATGATTAAGAAACCTGACTTGGCT 1472
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1473 GAAGACACAAACATTTGGGTCCACCATCTTAACCTCAAGGCACCTGATGAGACCA 1532
 Qy 481 PheThrGlySerSerLysIleLeuTyriIleIleLeuGlyAspSerGlnGlyArgLeu 500
 Db 1533 TTTTACTGGAGTTCTTAATTTCTGTATCATTTAAAGGAGACAGTGAAGGACGCTG 1592
 Qy 501 GlyValAspThrAspProHsiThrAsnThrGlyTyraIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTGACACAGATCCCATACCAACACCGAGATATGTCATTAATTAAGACCTCTTGAT 1652
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheValGluAsnProGluProLeuVal 540
 Db 1653 TTTGAAACAGACGCTGTTTCCACATGTTGTTCAAGACAGAAATCTGAGCCTCAAGTG 1712
 Qy 541 PheGlyValLysTyraAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1713 TTTGTGTGAGTACATGCAATGCAATCTTTTCCCAAGTTCCAGCTTATTGTGACAGATG 1772
 Qy 561 AsnGluAlaProGlnPheSerGlnIleValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGACACACCTCAATTTTCCCAACGATTTCCAAACGAAAGTCAGTGAAGATGTAGCT 1832
 Qy 581 TLeGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTy 600
 Db 1833 ATAGGCACTAAAGGGGCAAGTGAAGTCCAGAGAGTCCAGAGGTCGAGACATAGCAT 1892
 Qy 601 SerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyGluIlePhe 620
 Db 1893 TCACTGAGGGGAGACACAGAGGTTGGCTTAATTAATGACACGTCGATCGTGAATCTTT 1952
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyraArgValGlnValAlaThr 640
 Db 1953 AGTGTGCTCATTTGACAGAGAACCCGGAAGTCCATATCCGGTACCAAGTGTGTGGCCA 2012
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleMetAspVal 660
 Db 2013 GAAATAGGGGGGTCTTCTTGAAGCTGTGTGCAAGTTCACATGATCTTATGAGATGTG 2072
 Qy 661 AsnAspAsnProProAlaGlyLeuAlaLysAspTyraThrGlyLeuPhePheCysHisProLeu 680
 Db 2073 TATATACAACTCCCAAGCTAGCCAGAGACATACAGGGCTGTTCTTGTGCAATCCCTTC 2132
 Qy 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2133 AGTCACACTGGAAGTCTATTTGAGGCTACTGATGATGATCACTATTTATTCGGGGT 2192
 Qy 701 ProHsiPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2193 CCCCATTTTACATTTTCCCTGCGGACGTGGAAGCTTACAAACGATGGGAAGTTTCCAAA 2252
 Qy 721 TLeGlnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaValTy 740

Db 2253 ATCAATGTACTCATCTCCGCACTGTCTACAGGACACAGAGTTTGAGGAGAGGAGTAT 2312
 Qy 741 ValValLeuIleArgIleAsnAspGlyValArgProProLeuGlnGlyIleValSerLeu 760
 Db 2313 GTCTGCTTATCCGATCATATATATGAGGGGTGCGCCACCTTGAGAGCATTTGTTCTTA 2372
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2373 CCAATTACATTCGTCAGTTGTGTGAGAGAGAGATTTTCCGGCGACAGTCAACGACT 2432
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
 Db 2433 GGGATACCACTGTGGGCAATGAGTGGTATATCTGTCACCACTTCTGTGATTTGT 2492
 Qy 801 TLeIleLeuAlaValAlaPheIleArgIleValLysAspLysGlyLysAspAsnValGlu 820
 Db 2493 ATATATTTAGCAGTTGTGTTATCCCAATTAAGAAAGATTAAGGCAAGATTAATGTGA 2552
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2553 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAGAC 2588
 RESULT 10
 AX411163
 LOCUS Sequence 3810 from Patent WO0229103.
 DEFINITION AX411163
 ACCESSION AX411163
 VERSION AX411163.1 GI:21443868
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
 TITLE Gene expression profiles in liver cancer
 JOURNAL Patent: WO 0229103-A 3810 11-APR-2002;
 GENE LOGIC INC (US)
 FEATURES
 source location/Qualifiers
 1..3654
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="EMBL/Genbank Accession No. X83228"
 BASE COUNT 1052 a 804 c 777 g 1021 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1..99e-317 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.58% Indels: 0
 DB: 6 Gaps: 0
 US-10-025-380-1081 (1-832) x AX411163 (1-3654)
 Qy 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 93 ATGATCTTCAAGGCCATCTTCACTCCCTGTGTCTTATGCTTATTTAGGCACTGCA 152
 Qy 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysPheMetThrPheSerIleTyrglu 40
 Db 153 TATGCGCAAGAGGGGAAGTTTATGACACCCCTBAACCCATGACATTTCTATTATGAA 212
 Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe 60
 Db 213 GGCACGAAGACCGAGTCAATATATATCCAGTTTAAAGGCCATCTCTGTCGACTTTT 272
 Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluAlaArgGluGlyLeuLeuTy 80
 Db 273 GAACATACTGGGAGACAGCAACATATTTGTATGAACGGAGGAGACTTCTGTATTAC 332
 Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaLeuAsp 100

Db 333 AACAGAGCTTGGACAGGGAAACAAGATTACTCACAATCTCAGGTTGCACCTCGAC 392

Qy 101 ALaasngiYleIleValGluGlyProValProIlethriIleGluValIysAspIleAsn 120

Db 393 GCTAATGAAATATTATGATGGAGGGTCCAGTCCCTATCAACATTAAGTAAGGACATCAAC 452

Qy 121 AspAsnArgProThrPheIeuGlnSerIlyrGluGlySerValArgGlnAsnSerArg 140

Db 453 GACAAATCGACCCACAGCTTCTCCAGTCAAAAGTACGAAGGCTCAGTAGGACAGAACTCGC 512

Qy 141 ProGlyLysProPheLeuThrValIleAlaThrAspIleAspAspProAlaThrProAsn 160

Db 513 CCAGAAAGCCCTCTCTGTATGTCAATGCAACAGACTCGAATGAATCCGGCACTCCAAAT 572

Qy 161 GlyGlnLeuIlyrTyrgIlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrgPhe 180

Db 573 GGCCAGCTTTATTTACAGATTGTCAATCCAGCTTCCCAAGATCAACAATGTCAATGACTTT 632

Qy 181 GlnIleAsnAsnIlyrThrGlyValIleSerIleuThrArgGluGlySerGlnGlnLeuAsn 200

Db 633 CAGATCAACAACAACAAACCGGAGCCACTCTCTTATCCCAAGAGGATCTCAGGAATTGAAAT 692

Qy 201 ProAlaLysAsnProSerTyrgAsnLeuValIleSerValIysAspMetGlyGlyGlnSer 220

Db 693 CCTGTGAAGAACTCTCTATATGTGGAACTCAGTAAGTAAGACATGGAGGCCAGAGT 752

Qy 221 GlnAsnSerPheSerAspThrThrSerValAspIleIleValThrGlnAsnIleTyrgLys 240

Db 753 GAGAAATTCCTTACAGTATPCCACACTGTGGATATCATATGACACAGAAATTTTGGAAA 812

Qy 241 AlaProLysProValGluMetValGlnAsnSerThrAspProHisProIleLysIleThr 260

Db 813 GCACCAAAACCTGTGAGATGGTGGAAAACTCACTGATCTCACCCTCAAAATCACT 872

Qy 261 GlnValAlaGTPAsnAspProGlyValagIlyrTyrgLeuValAspLysGluLysLeuPro 280

Db 873 CAGGTCGGTGAAGATCCCGGTCCACAATATTCCTTAAGTTGAACAAAGAGAGCTGCCA 932

Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrgValThrGlnProLeuAspArg 300

Db 933 AGATTTCCCATTTTCAATTTGACCGAAGAAAGATATTACGTACACACCCCTGGACCGA 992

Qy 301 GlnGluLysAspAlaTyrgValPheTyrgAlaValAlaLysAspGluTyrgIlySerProLeu 320

Db 993 GAAGAAAAGAAAGCATATGTTTTTATCCAGTTCGAAAGATGAGTACGAAAAACACTT 1052

Qy 321 SerTyrgProLeuGlnIleHisValIysValIysAspIleAsnAspAsnProProThrCys 340

Db 1053 TCAATTCGCGTGAATTCATGTAAAGTAAAGATTAATGATTAATCCACTACATCACT 1112

Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluValGluGlyIysAsnSerIleGly 360

Db 1113 CCGTACCAAGTACCGGTAATTTGAGGTCCAGAGGAATGAAACATGGGTAAACAGATTCGG 1172

Qy 361 ThrIleuThrAlaHisAspArgAspGluGlnAsnThrAlaAsnSerPheLeuAsnTyrgArg 380

Db 1173 ACCCTTACTGCACATGACAGAGATGAABAAAATACTGCCAACAGTTTTCTTAACTACACG 1232

Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrgAla 400

Db 1233 ATTTGGAGCAAACTCCCAAACTTCCATGATGATGACTCTTCTATCAACAACTATCT 1292

Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerIleuLysGlnAspThrProGlnTyrgAsnLeu 420

Db 1293 GGAATGTTACAGTTAGCTTAAACAGTCCCTGAAGAAAGCAAGATATCTCCACATCAACTTA 1352

Qy 421 ThrIleGlnValSerAspLysAspPheTyrgThrLeuCysPheValGlnIleAsnValIle 440

Db 1353 ACGATAGAGGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAAAATCAAGTTATTT 1412

Qy 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrgIlyAsnLeuThrLeuAla 460

Db	1413	GATATCAATGATCAGACCCCACTCTTTGAAAATAACATGATTATGGAAACCTGACTCTTGGCT	14172
QY	461	GIuAspThrAsnIleGlySerThrIleuLeuThrIleGlnIalThrAspAlaAspGluPro	480
Db	1473	GAAGACACAAACATTGGGTCCACCATCTTTAACATCCAGCGCCACTGATGCTGATGAGCCA	15322
QY	481	PhetHrGlySerSerLeuIleLeuTYHisIleIlelysgIyaAspSerGluGlyArgLeu	500
Db	1533	TTTACTGGAGACTTCAAAATCTGATCATCTCAATAAAGGAGACAGTGAAGGAGCGCTCG	15922
QY	501	GIyValaAspThrAspProHisThrAsnThrGlyTYrValIleIleIleValysProLeuAsp	520
Db	1553	GGGGTGTGACAGATCCCCATPACCAACCGGGAATATGTCAATTAATTAAGAACCCCTCTTGAT	16522
QY	521	PhelGluThrAlaAlaValSerAsnIleValPheIysAlaGluAsnProGluProLeuVal	540
Db	1653	TTTGAAACAGACGCTGTTTCCACATCTGTCTTCAACAGGAAATCTCGACCTCTAGTG	17122
QY	541	PhelGlyValIlyTYrAsnAlaSerSerPheAlaIysPheThrLeuIleValThrAspVal	560
Db	1713	TTTTGTGTGAAGTCAATATGCAAGATTCTTTTSCCAAGTTCAAGCTTATGTGTACAGATGTG	17722
QY	561	AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIysValSerGluAspValAla	580
Db	1773	AATGAAGACCTCAATTTTCCCAACAGTAATTCACAGGAAAGTCAAGAGAGATGTAGCT	18322
QY	581	IleGlyThrIysValGlyAsnValThrAlaIysAspProGluGlyLeuAspIleSerTYr	600
Db	1833	ATTAGGCACTAAAGTGGGCAATGTATCTGCCAAGATCCAGAAAGTCTGGACATTAAGCTAT	18922
QY	601	SerLeuArgGlyAspThrArgIlyTYrPheIysIleAspHisValThrGlyIleIlePhe	620
Db	1893	TCACTGAGGGGAGACACAAAGAGTGTGGCTTAAATTTGCCACGATGACTGGTGAATCTTT	19522
QY	621	SerValAlaProLeuAspArgGluAlaGlySerProTYrArgValGlnValValAlaThr	640
Db	1953	AGTGTGGCTCATTTGACAGAGAAAGCCGGAAGTCCATATCGGATACAAAGTGGTGGCACA	20122
QY	641	GluValAlaGlyIysSerSerLeuSerSerValSerGluPheHisIleuIleuMetAspVal	660
Db	2013	GAAGTAGGGGGGTCTTCTTGAGCTCTGTGTCAAGTTTCAACCTGATCTTATGATGTG	20722
QY	661	AsnAspAsnProProArgLeuAlaIysAspTYrThrGlyLeuPhePheCysHisProLeu	680
Db	2073	AATGACAAACCTCCCAAGCTTAGCCAAAGACTACACGGGCTTGTCTTCTGCCATCCCTC	21322
QY	681	SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisIleuPheArgGly	700
Db	2133	AGTCCACCTGGAAATCATTTTTCGAGGCTACTGATGATCATGACCTTAATTTCCGGGGT	21922
QY	701	ProHisPheThrPheSerLeuGlySerGlySerLeuGlnIleAsnAspTYrPglValSerIys	720
Db	2193	CCCCATTTCATTTTCCCTCGGAGTGGAAAGCTTCAAAAACGACTGGGAAGTTTCCAAA	22522
QY	721	IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTYr	740
Db	2253	ATCAATAGTACTCATGCCCACTCTCTAACAGGACACACAGATTGAGAGAGGAATAT	23122
QY	741	ValValLeuIleArgIleAsnAspGlyIlyArgProProLeuGlnGlyIleValSerLeu	760
Db	2313	GTCTCTTTGATCCGCATCAATGATGGGGGTGGCCACCTTGGAAAGCAATTGTCTTTA	23722
QY	761	ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr	780
Db	2373	CCAGTTACATTCTGCAATGTGTGTGAAGAAAGTTGTTCCGGCCACGACGATCACACACT	24322
QY	781	GlyIleProThrValGlyMetAlaValAlaGlyIleLeuLeuThrThrLeuValIleGly	800
Db	2433	GGGATACCCACTGTGGGCAATGGCAATGGTATGCTGTGACACACCTTCTGGGATTTGGT	24922
QY	801	IleIleLeuAlaValAlaPheIleArgIleIysIlyAspIysAspIysIysAspAsnValGlu	820
Db	2493	ATATATTTTACAGATGTTGTTATCCGATATGAAGAGATTAAGGCAAAAGTATATGTTGAA	25522

QY 821 Ser1aGln1aSerGluValLeuProLeuArgSer 832
 DB 2553 AGTGTCAAGCATCTGAAGTCAACCTCTGAGAAAGC 2588

RESULT 11

HSRNALICA HSRNALICA 3654 bp mRNA linear PRI 01-JUN-1995
 LOCUS HSRNALICA H sapiens mRNA for LI-cadherin.
 DEFINITION X83228
 ACCESSION X83228.1 GI:854174
 VERSION X83228.1 GI:854174
 KEYWORDS LI-cadherin.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3654)
 AUTHORS Boettlinger, A., Kretz, B., Fieger, C., Dlouhy, B., Berndorff, D.,
 Goesener, R., and Tauber, R.

TITLE Molecular cloning of human LI-cadherin: evidence for a novel type of
 cadherin within the cadherin superfamily

JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3654)

AUTHORS Boettlinger, A.M.
 TITLE Direct Submission
 JOURNAL Submitted (05-DEC-1994) A.M. Boettlinger, Institut fuer Klinische
 Chemie & Bioch., Universitaetsklinikum Rudolf Virchow, Freie
 Universitaet Berlin, Spandauer Damm 130, 14050 Berlin, FRG

COMMENT Sequence overlapping with that under the accession number U07969.
 FEATURES location/Qualifiers

FEATURES

1..3654

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_id="uni-ZAP TM vector"

93..2591

/function="cell adhesion molecule"

/codon_start=1

/product="LI-cadherin"

/protein_id="CAA58231.1"

/db_xref="GI:854175"

/db_xref="SPTREMBL:Q15336"

/translation="MILQAHLSLCLMLYLATNGVQEGKFSGLKPMPTSIYEGGP
 SOTIFORKNPAPVATPELTGETDNI FVIEBEGLLYRALDRERSTRHVALADN
 GIVGEPVPTIKDINDRNPFLQSGVSGVSRKNSRPPILYVNAATLDDAPFN
 GQLYVQIVQPLMNNMNTYQINKTKGALISITRESGELNPAKPSINLVISYKDMG
 QSENSFSDTTSVDLITVENIWKAPKPEWENSTDPPIKITOVNMDPAQSLVAK
 EKLPRFPFSDIDEDIVTQPLDRBEDAVYFAVAADDEGKPLSYPLELHVAKDIN
 DNPTCPSPVTFEFOENRIGNSIGTLTAHDBENTANSFLNRYIVEQTPKLMG
 LFIOTYAGMLQAKOSLKKODTPQYNLTTEVSDKDPKTCFQVINDIDNTPRE
 KSDYGNITLADNTNIGSTLITQATDADPEPTGSKLHYTIKDSRGLGVDNPPPT
 NTGVYITIKPLDFTFAVSNITVFAENPEPLVGVKTNASSFAKFTLIVDNEAPFP
 SQHVQAKSEEDVAIGTKVGNVAKDDEGLDISLKGDTKGLKIDHVGELTSVAP
 LDRAGSPRYOVAVTEVGSLSVSEFHLIMDVANDNPRKLAKDVTGCFEFCFLSA
 PGSILFATDDOHLFRGPHFTFSLSGSLQNDWEVSKINGTHARLSRHEFEREY
 VVLRINDGGRPLEGISLPTFCSCVCEGSCFRPAGHQGIPVMAVAGLLTTLIV
 IGLTAVTFIRIKDKGNVESQASEVYKRS"

BASE COUNT 1052 a 804 c 777 g 1021 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.99e-317 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.58% Indels: 0
 DB: 9 Gaps: 0

US-10-025-380-1081 (1-832) x HSRNALICA (1-3654)

QY 1 MetIleuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 DB 93 ATGATACCTTCAGGCCCATCTTCACCTCCGTGTCTTCTTATGCTTATTTGGCAACTGGA 152

QY 21 TYRGIYGLNGIUGLYLYSPHESERGIYPROLEUYPROMETHRPHESEITETRYGIN 40
 DB 153 TATGCCAAGAGGGAGAGTTTATGTGACCCCTGAAACCATGACATTTCTATTATGAA 212
 QY 41 GYINGIUPROSERGINTLEIPEHEGINPELYALASNPProAlaValThrPhe 60
 DB 213 GGCCAGAAACCGGATCAATTAATATTCAGTTTAAAGCCAAATCTCTCTGTGACTTT 272
 QY 61 GILUETHRGIGLUINHRSPASNPILPEHEVALILEGLUARGGLUGLYLEUETRYTYR 80
 DB 273 GAACCTACTGGGGAGACAGCAACATATTTGTATGAAAGGGAGGACTTCTGTATTAC 332
 QY 81 ASNARGALALEUSPARGIUTHRARGSERTHIASNPLeuGlnValAlaLeuASP 100
 DB 333 ANAGAGCCTTGACAGGGAACAAATCTACTCAACATCTCAGGTGAGCCCTGGAC 392
 QY 101 ALAASNGIYLEILEVALIGLUYPROVALPROILETHRIEGLUVALYSAPPILEAN 120
 DB 393 GCTAATGAAATTATATGTGGAGGCTCCAGTCCCATACCATTAAGTGAAGACATCAG 452
 QY 121 ASPASNARGPROTHRPEHEUGINSELYTYRGLUGLYSERVALARGGIANSEARAY 140
 DB 453 GACAAATGACCCACGTTTCTCCAGTCAAGTACGAAGGCTCACTAAGCGAAGCTCCG 512
 QY 141 PROGLYLYSPROPHETUITYRVALAENALATHRASPLeuASPProAlaThrProASN 160
 DB 513 CAGAGAAAGCCCTTCTGTATGTCAATGACAGACATGATGATCCGGCACTCCCAAT 572
 QY 161 GYGLINLEUITYRTRGINTLEVALILEGINLEUPROMETILLEASNPValMetTYRPE 180
 DB 573 GGCCAGCTTATATACAGATGTCATCAGCTTCCATGATCAATGATGATGATGATCTT 632
 QY 181 GINTLEASNPANLSEYTHRGIALILESEITETHARGGLUGLYSERGINGLUENASN 200
 DB 633 CAGATCAACAAACAAAGGAGCATCTCTTACCCGAGAGGATCTTCAGAAATGAAAT 652
 QY 201 PROALYSEANPROSEYTHRAENLEUVALILESEYVALYSAPMETGYGINSEY 220
 DB 693 CTGCTNAGAACTCTTCTATATCTGATGATCTCAGTGAAGACATGGAGCCAGAGT 752
 QY 221 GLUASNPSEYSEYAPTHRTHSERVALASPILEILEVALTHRGLUENILETPLYS 240
 DB 753 GAAATTCCTTCAGTATACCAATCTGTGATATCATGATGATGATGATGATGATGATA 812
 QY 241 ALAPROLYSEYVALIGLWETVALIGUASNPSEYTHASNPProIleYsILETH 260
 DB 813 GCACCAAAACCTGTGAGAGTGGAGAAACCTCAACGATCTCAACCCATCAAAATCCT 872
 QY 261 GINVALARGTRPASNAPPProGlyYALAGINTYSEYVALASPIYSGIULYSLEUPRO 280
 DB 873 CAGGTGCGGTGAAATGATCCCGGTGACAAATATCTCTTATGATCAAAAGAAAGCTGCA 932
 QY 281 ARGPHETROPHESERILEAPGINGIUGLYASPILETYRVALTHRGLNPROLEUASPARG 300
 DB 933 AGATTCCCATTTCAATTGACCGAAGAGATATTACCTGACTCAGCCCTTGACACGA 992
 QY 301 GLUGLUYSEYSPALATRYALPHEYTRALAVALLYSAPGLUYTRGILYSPROLEU 320
 DB 993 GAAAGAAAGATCATATGTTTATTATGACGTTGCAAGATGATGATGATGATGATGAT 1052
 QY 321 SETYREPROLEUGLUILEHISVALYSVALYSAPPILEASNPASNProPOTHYR 340
 DB 1053 TCATATCCGTGGAATTTCTATGAAAGTTAAAGATATTATATGATATCAACCTACATGT 1112
 QY 341 PROSEYPROVALTHRVALPHEGLUVALINGLUASNGIULARGLEUGLYANSEYILEGLY 360
 DB 1113 CCGTCAACAGTAACCGATTTGAGTCCAGAGAAATGACAGCTGGGTAAACAGTATCGGG 1172
 QY 361 THRLEUTHRALHISAPARGASPGLUGIULASNPTRHALASNPSEYTHLEUANTYRARG 380
 DB 1173 ACCCTTACTGACATGACAGAGATGAAGAAATATCTGCCAAGTCTTCAAACTACAGG 1232
 QY 381 ILEVALIGLUINHRPProLYSPROMETASPGIYLeuPHEUULEGINTHRTRAL 400

Db 1233 ATGTGGAGCAAACTCCCAACCTTCATGATGAGCTCTTCCATACCAAACTTAAGCT 1292
 Qy 401 GlyMetLeuGlnLeuAlaValGlnSerLeuValGlnAspThrProGlnIleAsnLeu 420
 Db 1293 GGAATGTTACGTTAGCTTAACAAGTCTTGAAGAAGCAAGTACTCTCCAGTACACTTA 1352
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuGlnValGlnIleAsnValIle 440
 Db 1353 ACGATAGAGGTGTCTGACAAAGATTTCAGACCTTTGTTTGTGCAAAATCAAGTTATT 1412
 Qy 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlnAsnLeuThrLeuAla 460
 Db 1413 GATATTCATATGATGACGACCCCATCTTTGAAATAATCAATTAAGTAAACCTGACTTGTCT 1472
 Qy 461 GluAspThrAsnIleGlnSerThrIleLeuThrIleGlnIleAsnAspAlaAspGlnPro 480
 Db 1473 GAAACACAAACATGGGTCTCCATCTTAACATCAAGCCACTGATGCTGATGAGCCA 1532
 Qy 481 PheThrGlnSerSerLysIleLeuTyrHisIleIleLysGlnAspSerGlnGluYArgLeu 500
 Db 1533 TTTTACTGGAGTTCTAAATTTCTGTATCATATCATTAAGGAGACAGTGAAGGACGCTG 1592
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTTCACACAGATCCCATACCAACACCGGATATGTCATTAATTAAGCCCTCTGAT 1652
 Qy 521 PheGlnThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuVal 540
 Db 1653 TTTGAACACAGCTGTTCCAAATGTTGTTCAACACAGAAATCTTGAGCCCTTAAGTG 1712
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1713 TTTGGTGTGAAGTACATGACAGTTCTTTGGCCAGTTCAAGCTATTTGACAGAGTGTG 1772
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGAAGCACCTCAATTTTCCAAACGATATTCACACGAAAGTCAAGTGAAGTGTAGCT 1832
 Qy 581 IleGlnThrLysValGlnAsnValThrAlaLysAspProGluLysLeuAspIleSerTyr 600
 Db 1833 ATAGGACCTAAAGTGGCAATGATGCTCCAGAGATCCAAAGGCTCGACATTAAGCTAT 1892
 Qy 601 SerLeuArgLysAspThrArgGlyTyrLeuLysIleAspHisValThrGlyLysIlePhe 620
 Db 1893 TCCTGAGGGGAGACACAAAGGTGGCTTAAATTCACACGCTGAGTGTAGATTTT 1952
 Qy 621 SerValAlaProLeuAspArgLysIleGlySerProTyrArgValGlnValAlaThr 640
 Db 1953 AGTGTGCTCATTTGACAGAAAGCCGAAAGTCCATATCGGGTACAAAGTGTGGCCACA 2012
 Qy 641 GluValGlyGlnSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2013 GAAGTATGGGGGGTCTTCTTCTGAGCTCTGTCTGAGAGTTCACCTGATCTTATGATGTG 2072
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2073 AATGACAAACCTCCAGGCTAGCCAGGACTTAACAGGAGCTTGTCTTCTCCATCCCCCTC 2132
 Qy 681 SerAlaProGlnSerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2133 AGTGCACCTGAGAGTCTCATTTTCAGAGGCTACGATGATCAGACCTTAATTTCCGGGCT 2192
 Qy 701 ProHisPheThrPheSerLeuGlnGlySerGlnIleAsnAspThrProGluValAspLys 720
 Db 2193 CCCCATTTTACATTTTCCCTGCGCAGTGAAGCTTCAAAAACACACTGGGAAAGTTTCCAA 2252
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluAlaTyr 740
 Db 2253 ATCAATGATCTATGCTGACCTGTCTACAGGACACAGAGTTTGAAGAGAGGAGTAT 2312
 Qy 741 ValValLeuIleArgIleAsnAspGlyArgProProLeuGlnGlyIleValSerLeu 760

Db 2313 GTGCTTGTATCCGATTCATGATGAGGGGTGCGCACCTTGGAGGCAATGTTCTTTA 2372
 Qy 761 ProValThrPheCysSerCysValGlnLysSerCysPheAspProAlaGlnIleThr 780
 Db 2373 CCAATTACATTTCTGCGAGTTGTGTGAAGAGTGTGTCCGCGCAGAGGTCCACGACT 2432
 Qy 781 GlyIleProThrValGlnMetAlaValGlnIleLeuLeuThrThrLeuValIleGly 800
 Db 2433 GGGATACCACTGTCGGCAGTGGCATGCGATGATCTGACCAACCTTCTGTGATTTGT 2492
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGln 820
 Db 2493 ATATTTTTCAGCTGTTGTTATCCCATTAAGAGATTAAGCAAGATTAATGTTGAA 2552
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2553 AGTCTCAAGCATCTGAATCAAACTCTGAGAGC 2588
 RESULT 12
 AF177669 3351 bp mRNA linear ROD 05-JUN-2001
 LOCUS AF177669
 DEFINITION Mus musculus liver-intestine cadherin mRNA, complete cds.
 ACCESSION AF177669
 VERSION AF177669.1 GI:5762489
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS Angres,B., Kim,L., Jung,R., Gessner,R. and Tauber,R.
 TITLE LI-cadherin gene expression during mouse intestinal development
 JOURNAL Dev. Dyn. 221 (2), 182-193 (2001)
 MEDLINE 21270055
 PUBMED 11376485
 REFERENCE
 2 (bases 1 to 3351)
 AUTHORS Angres,B., Kim,L. and Tauber,R.
 DIRECT SUBMISSION
 JOURNAL Submitted (14-AUG-1999) Institut fuer Klinische Chemie und Biochemie, Virchow-Klinikum der Humboldt-Universitaet zu Berlin, Augustenburger Platz 1, Berlin 13353, Germany
 LOCATION/Qualifiers
 FEATURES
 source
 1..3351
 /organism="Mus musculus"
 /strain="Swiss Webster/NIH"
 /db_xref="taxon:10090"
 /cissue_type="intestine"
 /dev_stage="17 day embryo"
 44..2527
 /note="LI-cadherin; cell-cell adhesion molecule"
 /codon_start=1
 /product="liver-intestine cadherin"
 /protein_id="AAD51125.1"
 /db_xref="GI:5762490"
 /translation="MWSAQLHFLCLTLVLCGYSGEKGFSGLKMPMSIFEGEBS
 OYRQKTPNPATVETELCETGDIKIEKMDGLHYTRALDRPTAVVHHQLALDSHG
 AIVDEVPITTEVKDINDRPFLLSKIEGSTRQNSRPEKPMYNADLDLPATPG
 QLEQIVQLPQINDVMYFQIDSKALSLTEGSEQLDPVNSYNLVSXVDMGQ
 SENSFSDTYVDISIRENIWKAPBEVIRENSTDPHPKITOVQNNNDPAQYSLVKE
 KLSPPFSIDSGNLYVQALDREKNSHVPATADENGKPLAYPELEYVAVIIND
 NPCTGSPVTEVOENELGNSIGIFEAHMDANNINSLIKYLVDQTPVPSGL
 FLIGVEGVQVLSKQSLKQDSPOYNLSIEVSDNDPKTCYQNVVIDINDQPIFET
 SNYKTSLESDRTAISTIIITATNDADEFITSSKILKIVAGDIBGRLEVTDPTTN
 AGYVKKIKRPLDPTSPSIVSQANPELVKGIYVNSFASFSFLIVDVAEVPF
 QRIFOANVSQAAVAGSRVGNVARDPBGITVYSLSLGNRGMKIDSVGEI.FSAPL
 DRETSYVRVQVAVATEVGSSLSSTADFLVLYVDVNDPRLAKDYTGIF.FCHPSAP
 GSLFEVDDDOQS.LRPEFTPALRGEGQSDMEVSKINGTARLSTHRTREDOVN
 IPIRINDGQPMNEGTVLPVTFQCCVGESSCRPRAGRDGIFTVGMAGITLLTFLVI
 GIILAVPIRKMDVRENPPSPKRPAS"
 BASE COUNT 966 a 748 c 746 g 891 t
 ORIGIN
 Alignment Scores:

Pred. No.: 1,33e-252 Length: 3351
 Score: 3450.00 Matches: 658
 Percent Similarity: 88.21% Conservative: 75
 Best Local Similarity: 79.18% Mismatches: 94
 Query Match: 79.84% Indels: 4
 Gaps: 1

US-10-025-380-1081 (1-832) x AF177669 (1-3351)

Db 2 IleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGlyTyr 21
 44 ATGATGTCGCCCGAGCTTCACTTCTGCTGCTTCTCACTTATTTAGACCTGGAGAT 103
 Qy 22 GlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlnGly 41
 Db 104 GGGGAAAGAGGAAAGTTCCAGCGGTCCCTTAAGCCCAAGACATTTCTCATTTTGAAGGC 163
 Qy 42 GlnGluProSerGlnIleIlePheGlnPheGlyAlaAspProAlaValThrPheGlu 61
 Db 164 CAAGAACCGAGTCAAGTCATTTCTCACTTTTAAGCAACCTCCAGCTGAGCTTGCAG 223
 Qy 62 LeuThrGlyLysThrAspAsnIlePheValIleGluArgGlnGlyLeuLeuTyrTyrAsn 81
 Db 224 CTAAACGGAGAGCAGATGATGATTTAAGATGAAGAAAGATGAGACTTGTATCACACA 283
 Qy 82 ArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAspAla 101
 Db 284 AGAGCCCTGGACAGAGAAACAGAGCGGTTCATCATCTGCAGCTTGAAGCTTGGATTC 343
 Qy 102 AsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsnAsp 121
 Db 344 CATGAGCTTATGAGGAGCGGTCCAGTCCCATCCATAGAGTCAAGAGCATCAATGAC 403
 Qy 122 AsnArgProThrPheLeuGlnSerLysTyrGlnGlySerValArgGlnAsnSerArgPro 141
 Db 404 AACGACCCAGCGTTCTCCAGTCAAAATATGAAAGCTCAGAGGCAAGACTCTCGGCCA 463
 Qy 142 GlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspArgProAlaThrProAsnGly 161
 Db 464 GGAAGGCTTTCAGATGATGATCAATGATCAAGATCTGATGATCCGAGTACCTCCAAAGGC 523
 Qy 162 GlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPheGln 181
 Db 524 CAGCTTTTATTCAAATTTGTCATCCAACTTCCCAATCAAGATTTATGATCTTTCA 583
 Qy 182 IleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlnGlySerGlnGlnLeuAsnPro 201
 Db 584 ATCGACAGCAAAACAGGGGCAATATCGCTTACCCAGAGAGATCCCAAGATTTGATCCA 643
 Qy 202 AlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSerGlu 221
 Db 644 GTGAAGATCTTCTCTACACCTGTGTCTCGGTGAAGAGCATGGGGGCAAGATGAG 703
 Qy 222 AsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLysAla 241
 Db 704 AATTCCTTCACTGATCTACATATGATGATATTTCTATTAAGAGAAATTTCTGAAAGCA 763
 Qy 242 ProLysProValGlnMetValGluAsnSerThrAspProHisProIleLysIleThrGln 261
 Db 764 CCAGAACCCCGAGATTAAGAGAAACCTGACTGATCTCAACCCATCAAAATCATCTCAG 823
 Qy 262 ValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGlnLysLeuProArg 281
 Db 824 GTGAGTGTGAATGAGCCAGGGGCCAGATTTCTTAATGCAACAGAGAAACCTGTGCGCG 883
 Qy 282 PheProPheSerIleAspGlnGlnGlyAspIleTyrValThrGlnProLeuAspArgGlu 301
 Db 884 TTCCATTTCTGATCGACCAAGAGAAATTTTATGATCTCAGGCTTGGACCGGGAG 943
 Qy 302 GlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeuSer 321
 Db 944 GAAAGAACTCACTGATTTTCTTTTTCGCACTGCCAAGAGATGAGAAATGAGAAACCACTTGA 1003

Qy 322 TyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCysPro 341
 Db 1004 TATCCACTGGAATTTATTTGAAAGTATTTAGCATTAATGACACCAACCACTGCTG 1063
 Qy 342 SerProValThrValPheGlnValGlnGluAsnGluArgLeuGlyAsnSerIleGlyTyr 361
 Db 1064 TCTCCAGTACTGATTTGATTAAGTCCAGGAGAAATGAAACATTTGGTAAACAGCATCGGGATC 1123
 Qy 362 LeuThrAlaHisAspArgAspGlnGluAsnThrAlaAsnSerPheLeuAsnTyrArgIle 381
 Db 1124 TTGAGGCCCATGATATGAGACGAGCTTAACATCAATCAAGATTTTGAATTAACAGCTT 1183
 Qy 382 ValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAlaGly 401
 Db 1184 GTAGACCAAAACCCCAAGTTCCCTCAGATGAGCTTTTCTCATTTGGTGAATGAGGGA 1243
 Qy 402 MetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeuThr 421
 Db 1244 AAGGTTCAAGTTAAGTAAACAGTCTTGAAGAGCAAGACAGCTCTCAGATCAACTTAAGT 1303
 Qy 422 IleGlnValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIleAsp 441
 Db 1304 ATGAGAGTGTGACGTAATGATTTCAAGACTCTCTGTTATTTCAATGTCACGTTATGAT 1363
 Qy 442 IleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAlaGlu 461
 Db 1364 ATCAATGATCAAGTTCCATCTTTGAAACATCAATTAAGAAAGCAAGACTGTCTGAA 1423
 Qy 462 AspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPhe 481
 Db 1424 GACACCGCCATCGAGTCCACATCTTAATCATCAACATCAACATCTGATGAGCCCTT 1483
 Qy 482 ThrGlySerSerTyrIleLeuTyrHisIleIleLeuGlyAspSerGlnGlyArgLeuGly 501
 Db 1484 ACAGGAGGCTTAATAATCTGTAACAGATTTGACAGAGACACGAGGAGAACTATGAA 1543
 Qy 502 ValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAspPhe 521
 Db 1544 GTTGTCAAGATCCCAAGCAATGACAGATATGATCAAGATCAAAAGCCTTGAATTTT 1603
 Qy 522 GlnThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuValPhe 541
 Db 1604 GAAACTCAACCAAGTTGACAGCATTTGTTCCAAAGCGGAAATCTGAGCCACTGGTGA 1663
 Qy 542 GlyValLysThrAsnLysSerSerPheAlaLysPheThrLeuIleValThrAspValAsn 561
 Db 1664 GGCATGAGATGACATGATGATGATTTTGTGCTCTCTGAGCTGATTTGTGACGATGAT 1723
 Qy 562 GluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAlaIle 581
 Db 1724 GAAGTGTCTGTCTTCCCGGAGGAGATATTTCCAAAGCAAGCTCAGCGAGATGCTGTGTG 1783
 Qy 582 GlyThrLysValGlyAsnValThrAlaLysAspProGlnGlyLeuAspIleSerTyrSer 601
 Db 1784 GCGAGTGAAGTGGGCAACGTGACTGCCAGGAGATCTCGAAGCGCTGACTGAGATTAATCA 1843
 Qy 602 LeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGlnIlePheSer 621
 Db 1844 CTGAAGGCAATATGAGAGGCTGCTTAATATGACTCTCTCACTGCTGAGATTTTATG 1903
 Qy 622 ValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaAlaThrGlu 641
 Db 1904 GCTGCTCCGCTGATAGGAGAAAGAAAGTGTATTCGATTAACAAGTGGTGGCCACTGAA 1963
 Qy 642 ValGlyLysSerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspValAsn 661
 Db 1964 GTAGTGGGTCCTCTGAGATTTACCGCAGATTTCCACTGCTGCTCAAGATGTGAAT 2023
 Qy 662 AspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeuSer 681
 Db 2024 GACAAACCCCTGCTGCTGATGAAGACTCAACAGGCTTGTCTTGCACATCCCTCAGT 2083
 Qy 682 AlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGlyPro 701

```

|||||
Db      2084  GCGCCGCGAGCCCTTCCTTGAGTCACGACATCAACGACGATCATTCCGAGGCC 2143
|||
Qy      702  HiaPheThPheSerLeuGlySerGlySerLeuInaenAptPgiUvaSerLytle 721
|||
Db      2144  AAGTTTACATTCGCGCTTGAGAGAAAGCTTCAAAAGTGACTGGAGATTCAAAATC 2203
|||
Qy      722  AasnGlyThrhIsalaAglLeuSerThrhGhiThrhApePheGluGluAglAlyrVal 741
|||
Db      2204  AATGGTACACATGTGTAACTCTCCACGAGACACAGCTTTGAGAAACAAGTTATTAAC 2253
|||
Qy      742  ValLeuIleArgIleAenAapGlyGlyArgProProLeuGluGlyIleValSerLeuPro 761
|||
Db      2264  ATTCCAAATCCGCAATGATGATGAGGAGCAGCAACCCATGAGAGGAGCTGCTTCAACA 2323
|||
Qy      762  ValThrhPheCysSerCysValGluGlySerCysPheAhrProAaGlyIleGlnThrhGly 781
|||
Db      2324  GTTACTCTTGTCCAGCTGTGTGAAAGAAAGCTGTTCCGCGCAGCAGCAGAGATGG 2383
|||
Qy      782  IleProThrValGlyMetAlaValGlyIleLeuLeuThrhThrhLeuLeuValIleGlyIle 801
|||
Db      2384  ATACCCACTGTGGGCAATGAGGAGTTGATTAATCTTGAACACATTTCTGCTCATTTGATTA 2443
|||
Qy      802  IleLeuAlaValAlaPheIleArgIleLeuLysAapLysGlyLysAapAenValGlySer 821
|||
Db      2444  ATTTTAGCAGTGTCTTCATCCGCAAGAGAAAGATATAA-----GTTGAAAT 2491
|||
Qy      822  AlaGluAaSerGluValLysProLeuAhrSer 832
|||
Db      2492  CCGCAATGCCCTTGAATTAAGCCCTTGAAGAGC 2524
|||
RESULT 13
LOCUS      D87912          3414 bp      mRNA      linear      ROD 06-OCT-2000
DEFINITION Mus musculus mRNA for BHLH-cadherin, complete cds.
ACCESSION  D87912.1 GI:9392287
VERSION     D87912.1
KEYWORDS    Mus musculus (strain: BALB/c) foetal liver Prob cell cell_line:3889
SOURCE      cDNA to mRNA.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases)
AUTHORS     Ohnishi, K., Shimizu, T., Karasuyama, H. and Melchers, F.
TITLE       The identification of a nonclassical cadherin expressed during B
            cell development and its interaction with surrogate light chain
            J. Biol. Chem. 275 (40), 31134-31144 (2000)
JOURNAL     20469471
MEDLINE     2 (bases 1 to 3414)
REFERENCE   Ohnishi, K.
TITLE       Direct Submission
            Submitted (18-SEP-1996) Kazuo Ohnishi, National Institute of
            Health, Department of Immunology, 1-23-1, Toyama, Shinku-ku,
            Tokyo 162, Japan (E-mail: ohnishi@nih.go.jp,
            Tel: 03-5285-1111 (ex. 2133), Fax: 03-5285-1150)
FEATURES
source      Location/Qualifiers
            1..3414
            /organism="Mus musculus"
            /strain="BALB/c"
            /db_xref="taxon:10090"
            /cell_line="3889"
            /cell_type="Prob cell"
            /tissue_type="foetal liver"
            93..2576
            /codon_start=1
            /product="BHLH-cadherin"
            /protein_id="BAB03264.1"
            /db_xref="GI:9392288"
            /translation="MYSAQHLHFCILTYLTCTGYGEGKFSGLKPMWTSIFPGSPS
            QVLPQRKINPVAATFELTGETDGIKIKEDGLYHTRALDRFRAVHHQLALADSHG
            ALVDGPVLTIEVKDINDNPTFLQSKYBSVRQNSRPGKPFMYNAVALTIDDPATNG
            QLFYQIVQLFQINDVMYFQIDSKTGLSITPGSQELDPVKPNSYLVSVKMDMGQ

```

```

SENSFDDTVYVDSIRINIWKAPEVEIRENSTDPHPKITQVOAMDPGAQSLVWKE
KLSPPESIDSENIYVTOALDEEGKSHVFPFATADENKPLAYPELEIVKYLIND
NPTCLSPVTFEVOENEPNIGIFEAHMDMEANNINSILKYKLVDPQKPSDGL
FLIGVEKGQVQLSKSLKKODSQVNLSEIVSDVDRKTCYIOVNVITDIDQPIRET
SNYGSKTLSLSDIAIGSTILITQADDEPTGSSKLYIKVQSDTGRLEVDPTTN
AGYVKIKLPDPEFQVSSIVFOAENPEPLVKIEVNASFAPFELIVDNDVPPF
QRIFOANVSEDAVAGSVGNVTADEGLVSYSLGNMKGMLKIDSVGEISAPL
DRETEVYRQVAVGVSSLSSTADPHLVLTVDVNDNPRLAKDYTGFEPCPLSP
GSLIFETDDQOGLARPKPTALRGEGIOSDEVKINQTHARLSTHRPREOYVN
IPRINDGQPPHEGTFVLPVTECCQVBGSQCFRPAQRDSIPVGNVAILTLTFLVI
GITLAVFTIMRDXENPQSPENKPLRS"
BASE COUNT      967 a      766 c      774 g      907 t
ORIGIN
Alignment Scores:
Pred. No.:      1,36e-252      Length:      3414
Score:          3450.00      Matches:      658
Percent Similarity: 88.21%      Conservative: 75
Best Local Similarity: 79.18%      Mismatches: 94
Query Match:      79.84%      Indels:      4
DB:              10      Gaps:      1
US-10-025-380-1081 (1-832) x D87912 (1-3414)
Qy      2  IleGluAlaIahIsPheHIsSerLeuCyLeuLeuMetLeuThrhAlaThrhGlyTyr 21
|||
Db      93  ATGTCGTCTGCCAGCTTCATCTTCCTGTGCTTCTACACTTATTTGACCTGTGATAT 152
|||
Qy      22  GlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrhPheSerIleTyrGluGly 41
|||
Db      153  GCGGAAGAAGGGAAGTTCAAGCGGTCCCTGAAAGCCATGACATTCCTCATTTTGAAGGC 212
|||
Qy      42  GlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrhPheGlu 61
|||
Db      213  CAAGAACCAGTCAAGTCAATTCAGTTTAAACCAACCTCCACCTGTGACTTTCGAG 272
|||
Qy      62  LeuThrhGlyLuthrhAapenIlePheValIleGluAhrGluGlyLeuLeuTyrTyrAsn 81
|||
Db      273  CTAAACGGAGACACATGATGATATTTAAGATGAAAGATGACTTCTGTATCACACA 332
|||
Qy      82  ArgAlaLeuAapArgGluThrhArgSerThrhIsaSnLeuGluValAlaAlaLeuAapAla 101
|||
Db      333  AGAGCCCTGGACAGAGAAACAAGAGCGGTTCAATCATCTGAGCTTGACAGCTTGATTC 392
|||
Qy      102  AasnGlyIleIleValGluGlyProValProIleThrhIleGluValLysAapIleAenAap 121
|||
Db      393  CATGAGCTATAGTACGAGTCCAGTCCCATCAACCATCAAGATCAATCAATGAC 452
|||
Qy      122  AasnArgProThrhPheLeuGlnSerLysTyrGluGlySerValArgGlnAenSerArgPro 141
|||
Db      453  AACCGACCAAGCTTTCCTCCAGTCANAAATGAAAGCTCATGTAGGCAAGAACTTCGCCCA 512
|||
Qy      142  GlyLysProPheLeuTyrValAasnAlarhAapLeuAapProAlarhProAenGly 161
|||
Db      513  GGAAGACCTTTCATGATGACATGCTACAGATCGATGATCGGCTACTCCCAATGGC 572
|||
Qy      162  GlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAenAenValMetTyrPheGln 181
|||
Db      573  CAGCTTTTATTCAAATGTCTCATCCAACTCCCAATCAACCATGTTATGTACTTTCA 632
|||
Qy      182  IleAasnLysThrhGlyAlaIleSerLeuThrhArgGluGlySerGlnGluAenPro 201
|||
Db      633  ATCGAAGCAAAACAGGGGCAATATGCTTACCCCAAGAGATCCAGAAATTTGATTCAC 652
|||
Qy      202  AlaLysAapProSerTyrAenLeuValIleSerValLysAapMetGlyGlnSerGln 221
|||
Db      693  GTGAAGAAATCTCTCTCAACAACCTGCTCGGTGAAGAGACATGGGGGCGCAAGTGAAG 752
|||
Qy      222  AasnSerPheSerAapThrhThrhSerValAapIleIleValThrhGluAenIleThrpLysAla 241
|||
Db      753  AATTCCTTCAGGATCTCATATGATGATATTTCTATTAGAGAAATATCTCGAAAGCA 812
|||
Qy      242  ProLysProValGluMetValGluAenSerThrhAapProIleProIleLysIleThrhGln 261

```



```

Db      813  CCGAAGCCCGTGGAGATTAGAGAAAATCTGACATCTCTCAACCCCACTAAATCACTCAG 872
Qy      262  ValArgTpaAnaPpProGlyValAgIntYserLeuValAspLyGulYLeuProArg 281
Db      873  GTGACAGTGAATGACCCAGGGGCCAGATATCTTATGCAACAAGAGAAAGCTGTCCGCC 932
Qy      282  PheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArgGlu 301
Db      933  TTCCATCTTCGATGACCAAGAGAAATTTATGACTGACGCCCTTGACCCGAG 992
Qy      302  GluYAspAlaTyrValPheTyrAlaValAlaYAspGlnTyrGlyYLeuProLeuSer 321
Db      993  GAAAAGAACTCACTGTTTCTTCTTGCACTGCCAAGATGAGATGAAAAACCTTGCA 1052
Qy      322  TyrProLeuGlnIleHisValIleValIleAspIleAsnAspAsnProProThrCysPro 341
Db      1053  TATCCTCTGGAAATTTATGTAAGTTATGACATTAATGACAAACCCACGACATGTCTG 1112
Qy      342  SerProValThrValPheGlnValGlnGluAsnGlnLysGlnYAsnSerIleGlyThr 361
Db      1113  TCTCAGTGAATGATTTGAAATGTCAGAGAAATGACCATTTGGGTAAACAGCATGGAGATC 1172
Qy      362  LeuThrAlaHisAspArgAspGlnGluAsnThrAlaAsnSerPheLeuAsnTyrArgIle 381
Db      1173  TTTCAGGCCCATGATATGACAGAAAGCTAAACAACATCAACAGTATTTGAAATACAGCTT 1232
Qy      382  ValGlnGlnThrProYLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAlaGly 401
Db      1233  GTAGACCAAAACACCAAAAGTCCCTCAGATGACATTTTCTCATTTGGTGAATATGAGGGA 1292
Qy      402  MetLeuGlnLeuAlaYAspGlnSerLeuYLeuGlnAspThrProGlnTyrAsnLeuThr 421
Db      1293  AAGGTTCAGTTAACTAAACAGTCTCTGAAAGAGACAGACAGTCTCTGATCACTTAAGT 1352
Qy      422  IleGlnValSerAspIleAspPheLeuPheLeuCysPheValGlnIleAsnValIleAsp 441
Db      1353  ATCAGAGTGTCTGACAGTATTTCAAGATCTCTGTTATATTCAGTCAAGCTTATTTGAT 1412
Qy      442  IleAsnAspGlnIleProIlePheGlnYLeuSerAspTyrGlyAsnLeuThrLeuAlaGlu 461
Db      1413  ATCATATGATCATGATCTCCATCTTTGAAACATCAATTAATGAAACAGACCTCTCTGAA 1472
Qy      462  AspThrAsnIleGlySerThrIleLeuThrIleGlnIleThrAspAlaAspGluProPhe 481
Db      1473  GACACAGCATCGATCCACATCTCTATCATTCAGATCAAGATGCTGATGAGCCCTTT 1532
Qy      482  ThrGlySerSerIleLeuTyrHisIleIleLeuGlyAspSerGlnYArgLeuGly 501
Db      1533  ACAGGAGACTTAAATCTCTGTACAGATTTGACAGGAGACAGGAGGAGAACTGAA 1592
Qy      502  ValAspThrAspProHisIleThrAsnThrGlyTyrValIleIleYLeuYProLeuAspPhe 521
Db      1593  GTTGTTCAGATCCACAGCAACCAATGCAAGATATGTCAAGATCAAAAAGCCCTTGATTTT 1652
Qy      522  GluThrAlaAlaValSerAsnIleValPheYAsnIleGluAsnProGlnProLeuValPhe 541
Db      1653  GAAACTCAACAGATTAGCAGATTTGTTCAGAGGAAAAATCCGACCCACTGTGAA 1712
Qy      542  GlyValYAsnAlaSerSerPheAlaYAspPheThrLeuIleValThrAspValAsn 561
Db      1713  GGCAATAGATCAATGCTAGTTCTTTGCTCTCCCTCGAGCTGATTTGACGAGATGAT 1772
Qy      562  GluAlaPpProGlnPheSerGlnHisValPheGlnAlaYAsnValSerGluAspValAlaIle 581
Db      1773  GAAAGTCTGTGTTCCCAAGCAATATTCAGAGCAACGTCAGAGAGATGCTGCTGTG 1832
Qy      582  GlyThrIleYValGlyAsnValThrAlaYAspPpProGlnYLeuAspIleSerTyrSer 601
Db      1833  GGCGATAGTGGGCAAGCTGATCGCAGGATCTCTGAAAGCTGATCTGATGATTAACA 1892
Qy      602  LeuArgGlyAspThrArgGlyTyrPheYAsnIleAspHisValThrGlyGlnIlePheSer 621

```

```

Db      1893  CTGAAGCAATATGAGAGGCTGTAAATGACTCTGTCACTGGTGGATATTTAGT 1952
Qy      622  ValAlaProLeuAspArgGlnAlaGlySerProTyrArgValGlnValAlaIleThrGlu 641
Db      1953  GCTGCTCCGCTGATAGGAAACAGAAAGTGTATGATGATCAAGAGTGGGCACTGAA 2012
Qy      642  ValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValAsn 661
Db      2013  GTAGGTGGGTCTCTGATGATTTACGGCAGATTTCACTGCTCTCAAGATGTGAAT 2072
Qy      662  AspAsnProProAlaGlnLeuAlaYAspIleTyrThrGlyLeuPhePheCysHisProLeuSer 681
Db      2073  GACAAACCCCTGCTGCTAGCTTAAGACTACACAGGCTTGTCTTCTCCATCCCTCAGT 2132
Qy      682  AlaProGlySerLeuIlePheGlnAlaIleThrAspAspAspGlnHisLeuPheArgGlyPro 701
Db      2133  GCCCTTGGAGGCTCATCTTTGAGTCACTGAGAGATGACCAAGTCACTTCGAGAGCCC 2192
Qy      702  HisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspIlePpGlnValSerIle 721
Db      2193  AAGTTTACGTTCGCCCTTGGAGAGAAAGCTTACAAAGTACCTGGGAAGTTCAAAATC 2252
Qy      722  AsnGlyThrHisAlaArgLeuSerThrArgHisIleThrAspPheGlnGluArgAlaTyrVal 741
Db      2253  AATGGACACATGCTGACTCTCCACAGGACACACACGCTTTGAGAAACAAGTTATMAC 2312
Qy      742  ValLeuIleArgIleIleAsnAspGlyYArgProProLeuGlnGlyIleValSerLeuPro 761
Db      2313  ATTTCATTCGATCAATGATGAGGGGAGCCACCAATGAGAGACTGTCTTTACCA 2372
Qy      762  ValThrPheCysSerCysValGlnGlySerCysPheAspProAlaGlyHisGlnThrGly 781
Db      2373  GTTACTTCTCGCAGTGTGTGGAAGAGCTGTTCCGGCAGACAGGACAGAGATGGG 2432
Qy      782  IleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIle 801
Db      2433  ATACCCACTGTGGGCGATGGGAGTGGTATCTTCTGCCACATTCCTGGGTCATGTGATA 2492
Qy      802  IleLeuAlaValAlaPheIleArgIleYLeuYAspIleGlyYAspAsnValGluSer 821
Db      2493  ATTTTACAGATTTCTTCATCCGATGAGAGAAAGATAAA-----GTTGAAAT 2540
Qy      822  AlaGlnAlaSerGlnValYAspProLeuArgSer 832
Db      2541  CCGAGTCCCTGAGAAATAGCCTCGAGAGG 2573

RESULT 14
RNCADHER 3628 bp mRNA linear ROD 13-SEP-1994
LOCUS R.norvegicus (Sprague Dawley) mRNA for cadherin.
DEFINITION
ACCESSION X78997.1 GI:505562
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus.
Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 3628)
Berndorff,D., Gessner,R., Kreft,B., Schnoy,N., Lajous-Petter,A.M.,
Loch,N., Reutter,W., Hortsch,M. and Tauber,R.
Liver-intestine cadherin: molecular cloning and characterization of
a novel Ca(2+)-dependent cell adhesion molecule expressed in liver
and intestine
JOURNAL J. Cell Biol. 125 (6), 1353-1369 (1994)
MEDLINE 9426966
PUBMED 8207063
REFERENCE 2 (bases 1 to 3628)
Tauber,R.
Direct Submission
Submitted (15-APR-1994) R. Tauber, Institut fuer klinische Chemie
und Biochemie, Universitaetsklinikum Rudolf Virchow, Freie Uni.
Berlin, Spandauer Damm 130, 14050 Berlin, FRG

```


Db	1928	522	1888	542	1948	562	2008	582	2068	602	2128	622	2188	642	2248	662	2308	682	2368	702	2428	722	2488	742	2548	762	2608	782	2668	2688	802	2728	822	2776
Db	GTGTCCACAGATCCCATGACCAATATACAGATACGTCAAGATCAGAAAGCCTTGATTTT	GIuThrAlaValSerAsnIleValPheIysAlaGluAsnProGluProLeuValPhe	GAACCCAGACCAAGCTTACGTACGATGTGTTCAAAACAGAAACCCCTGAGCCCTGTGAT	GIValIySerTyrAsnAlaSerSerPheAlaIysPheThrLeuIleValThrAspValAsn	GGCATTAAGTACAAATGCTGATGTTCTTTGGCTCCCTTCGAGCTGACCGGAGCGAATGTAAT	GIuAlaProGluPheSerGlnHisValPheGlnAlaIysValSerGluAspValAlaIle	GAAGTGCCTGATTTCCCGACGAATATATCCAAAGCCACGTTAGCGAGATACGTCTATA	GIYThrIysValGIyAsnValThrAlaIysAspProGluGlyLeuAspIleSerTyrSer	GGCATTAAGTGGGCACTGTGACGTGGCGGATCCGGAAGGCGTGAAGTGTGATTTATCA	LeuArgGlyAspThrArgGlyTyrPheLeuValIleAspHisValIleThrGlyGluIlePheSer	CTGAAGAACAATTAAGAGAGTGGCTTAAATATCACTCCGTCACCGGAGCAATATTTAGC	ValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThrGlu	ACTGTCCTCGCTGAGACAGGAAACAGAAAGTGTGATGAGATACAAAGTGTGGCCATCGAA	ValGIyGIySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValAsn	GTAGGTGGGTCTCTCTGATGTTTCAACATATTTCCACTGCGTCATGATGATGATGAT	AspAsnProProlGlyLeuAlaIysAspTyrThrGlyLeuPhePheCysHisProLeuSer	GACAAACCCCTCCGCTAGCTAAGACATCAACAGGCTTGTTCTTGTGCATCCCTTAGT	AlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGlyPro	GCCCCGTGGAGCCTCATATTGTAACCACTATGATGACCGACGTCAAGTTGGAGGCC	HisPheThrPheSerLeuGlySerGIySerLeuGlnAsnAspTyrGluValSerIysIle	AAGTTTACGTTTGCCCTTGGCAGAGAAAGCTTACAAGATGACGTGGAAAGTGTCCAAATC	AsnGlyThrHisAlaIysLeuSerThrArgHisThrAspPheGluGluIysAlaIysVal	AATGGCACACATGTAGACTCTCCACCAAGACACACACCTTGGAGAACAGTTTATGAC	ValLeuIleArgIleAsnAspGlyGIyIysArgProProlGluGluIleValSerLeuPro	ATCCCAATCTCATATGATGATGGGGGGGACGCCACCATGAGAAAGGCATTGTCTCTTATCA	ValThrPheCysSerCysValGIyGlySerCysPheArgProAlaGlyIleGlnThrGly	GTTACTTCTGCGACAGTGTGTGGAGCGAAGTTGTTTCGGGCACAGCAGCAACAAGTTGGG	IleProThrValGIyMetAlaValGIyIleLeuLeuThrThrLeuLeuValIleGlyIle	ATACCCACCGGTGGCATGAGAGTGGTATCTTGCACCAATTTCTGTGATTTGATTA	IleLeuAlaValValPheIleArgIleIysIysAspArgGlyIysAspAsnValGIySer	ATTTTAGGGGTGTATTCATCCGTAATGAGAAAGATTA-----GTTGAGAT	AlaGlnAlaSerGluValIysProLeuArgSer	CCACAGTCCCTCGAAGATTAAGCCTTGAAGAC	
Db	1928	522	1888	542	1948	562	2008	582	2068	602	2128	622	2188	642	2248	662	2308	682	2368	702	2428	722	2488	742	2548	762	2608	782	2668	2688	802	2728	822	2776
LOCUS	RATPRPT																																	
DEFINITION	Rattus norvegicus proton-driven peptide transporter mRNA, 3' end cDNA																																	
ACCESSION	L46874																																	

REFERENCE
AUTHORS
TITLE
JOURNAL
BIOCHEM
MEDLINE
PUBMED
FEATURES
SOURCE

ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 1733)

ERICKSON, R.H., GUN, J.R., LINDEGROM, M.M., MCKEAN, D. and KIM, Y.S.
Regional expression and dietary regulation of rat small intestinal peptide and amino acid transporter mRNAs
Biochem. Biophys. Res. Commun. 216 (1), 249-257 (1995)

9667558
7468096

location/Qualifiers
1. 1733
/organism="Rattus norvegicus"
/strain="Wistar"
/db_xref="taxon:10116"
/clone="rpt-1"
/class_type="small intestine"
<1. .>1733
/codon_start=1
/product="proton-driven peptide transporter"
/protein_id="AAC42077.1"
/db_xref="GI:1082043"
/translation="LTGETDGLKLEDDGLYHARVUDRETRAVHRLRLAALDSQAI
VDGVPILIEVDKINDRPTFLQTKYEGSVRSRPEKPRMYNATLDLPATFNGOL
FYQVLIQPKINVMYFQIDNKTGALSTPEGSQVLDPIKNPYNLVASVDMGGONE
NSFEDTSVDITVRENIMWKAPEVIRENTIDPHIKITOVQNNBPVAHYSKLNKEL
POPEFSIDQGNLYVQPLRREKDSHFPAFKADENGRFLAYLEIRVYIDINDRP
PTCSQVTFVPEQNEPTLSSGIFRANDDANNINILMYRLNODTPVPSDELFL
IDEGKAVQLGKSLKKQDSPQYRLATVEVDIDPKTCSQVAVIDINDQPIPERSD
YGSSTLSQDPAIGSTLIIGATDDEPTSSKILYKIVGDTGRLVLEVVDTPWING
YVKIRKRLDETEPEPVSILVFKANPEPLVNGIEYNASFSFELVAVNEVPVFPQO
IFQANVEDTAIGTKVGTATVPDEGLTVSYLKNRGMKIDSVGDIFFSTAPLDR
ETESVYVQVVA"

BASE COUNT 525 a 398 c 393 g 416 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 9.94e-168 Length: 1733
Score: 2328.00 Matches: 446
Percent Similarity: 86.83% Conservative: 55
Best Local Similarity: 77.30% Mismatch: 76
Query Match: 53.88% Indels: 0
DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x RATTPEPRT (1-1733)

QY 62 LeuThrgIyguThrraPaenIlePheValIIleGIuArgGIuLeuLeuTYTTha 81
Db 1 CTAACAGAGAGAGAGATGTAATCTTAAGATGAAAGAGATGATCTTTACACACA 60
QY 82 ArgAlaLeuAaPaArgGIuThrArgSerThiAsnLeuGlnValAlaAlaLeuAspAla 101
Db 61 AGAGTCTCGACAGAGAAACAAAGGGCGGTTCATATCTGCGGCTTGACGCTTGATCT 120
QY 102 AenGIyIleValIGuGIyProValProIleThrIleGIuValIyAspIleAsnAsp 121
Db 121 CAGGAGCTATAGTGGATGGCCAGTCCCATCATAGAAAGTCAGAGATCATAGAC 180
QY 122 AaAaArgProThrPheLeuGlnSerIyTyGIuGIySerValArgIlnAaSerArgPro 141
Db 181 AACGACCTACATTTCTCCACAGAAATATGAAGCTCAGTGAAGCAGAACTCGCCCA 240
QY 142 GIyIySPProPheLeuTYTValAsnAlaThraSPLeuAaPaProAlaThrProAsnGIy 161
Db 241 GGAAGGCTTTCAATGTCGTAAGTCACAGACCTGATGACCCGGCTACCTCCAAATGC 300
QY 162 GlnLeuTYTArgIlnIleValIIleGIuLeuProMetIleAaAenValMetTyPheGIln 181

```

Db      301 CAGCTTTTATCAAAATGTCATCCAGCTCCAAAAACAAATGTAATCTTCAG 360
Qy      182 11leasnAnlyThrnglyAlaIleSerLeuThraArgLuglySerGlnGluLeuAsnPro 201
Db      361 ATGACAAACAACAAAGGGGCAATCTCGCTTACCCCAAGAGTTCTCAGGTAATGATCCA 420
Qy      202 AlAlAsnAnProSerTyraSnLeuValIleSerValIlyAspMetGlyGlnSerGlu 221
Db      421 ATAAAGATCCTTACTACAACTGTGTGTCTCTGTGAAGGACATGGGGGGCAGAAATGAG 480
Qy      222 AsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTrrplyAla 241
Db      481 AATTCCTTCAGCGATACCAATCTGTGATATTACTCTCAGAGAAACATCTGAAAGCA 540
Qy      242 ProlyAspProValGluMetValGluAsnSerThraAspProHisProIleIleThrGln 261
Db      541 CCAGAGCTGTGAGATTCGAGAAACCTTAATCTGATCTCCACCAACAAATCAGCAG 600
Qy      262 ValArgTrpAsnAnpProGlyAlaGlnTyraSerLeuValAspLysGlnLysLeuProArg 281
Db      601 GTGAGTGAATGAACCAAGTGGCCCACTAATCTTATCAACAAGAGAAAGCTGCCACAG 660
Qy      282 PheProPheSerIleAspGlnGluGlyAspIleTyraValThrGlnProLeuAspArgGlu 301
Db      661 TTTCATTCCTCGATCGACCAAGAGAAATATTAACCTGACCTGACCTTGACCTGAG 720
Qy      302 GluLysAspAlaTyraValPheTyraValAlaAlaLysAspGluTyrglyLysProLeuSer 321
Db      721 GAAAAGACTCTCATGTGTTTCTTTGCGACGCGCAAGAGCAAGAAATGSAAAACCCCTTGCA 780
Qy      322 TyrProLeuGlnIleHisValLysValLysAspIleAsnAspAnProProthCysPro 341
Db      781 TATCACTGGAATTCGTGGAAGATTATTAATGACATTAAACGCAACCCAGACATGCTC 840
Qy      342 SerProValThrValPheGlnValGlnGluAsnGluArgLeuGlnLysSerIleGlyThr 361
Db      841 TCTCAAGTGAAGTATTAATGAAGTCCAGAGAAATGAATCTTGGTACTAGCATTTGGGATC 900
Qy      362 LeuThrAlaHisAspArgAspGlnGluAsnThrAlaAsnSerPheLeuAsnTyraArgIle 381
Db      901 TTTCGTCGCCACATATGATGATGAAGGCCAACAATATCAACAGTATTTTGATGTCAGACTT 960
Qy      382 ValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyraAlaGly 401
Db      961 GTAGACCAACACCAAGTATCCCTCAGATGAGACTCTTCTCATTTGACGAATATGAGAGA 1020
Qy      402 MetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyraSnLeuThr 421
Db      1021 AAGGTCCAGTTGGGTAAACGCTCTTGAAAGACAGACAGTCTCAGTACCACTTAAACA 1080
Qy      422 IleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIleAsp 441
Db      1081 GTAGAGGTGCTGACATAGATTCAGACTCTGTTCTCTTCAAGTCAACGTTATTTGAT 1140
Qy      442 11leasnAspGlnIleProIlePheGlnLysSerAspTyrglyAsnLeuThrLeuAlaGlu 461
Db      1141 ATCAACGATCAGATTCCTCATCTTGAAGATCAGACTATGGAAGCAAGACTCTTCTCGAA 1200
Qy      462 AspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGlnProPhe 481
Db      1201 GACACAGCCCATCGATCCACATCTTAATCATCTCAAGCTACTGATGATGAGCCCTTT 1260
Qy      482 ThrGlySerSerLysIleLeuTyraHisIleIleLysGlyAspSerGlnGlyArgLeuGly 501
Db      1261 ACCGGAGCTCTTAATTCGTACAAAGATGTACAGGGAGACACTGAGGAGAGACTGGAA 1320
Qy      502 ValAspThrAspProHisThrAsnThrGlyTyraValIleIleLysLysProLeuAspPhe 521
Db      1321 GTTGTCACAGATCCCATGACCAATACAGGATACGTCAAGATCAGAAAGCCTCTTGATTTT 1380
Qy      522 GluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuValPhe 541
Db      1381 GAACCGAACCAAGTACTAGCATTTGTTCAAAGCAGAAAAACCTGAGCGCTGTGTAAT 1440

```

```

Qy      542 GlyValIleTyraSnAlaSerSerPheAlaLysPheThrLeuIleValThrAspValAsn 561
Db      1441 GGCATAGAGTACAAATGCTAGTCTTTGCTCTCTTGAAGCTGACCGTGAAGATGTGAAT 1500
Qy      562 GluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAlaIle 581
Db      1501 GAAGTCCGTGATTTTCCCGCAAAATATTCCAAGCCAAAGCTTATGCGAGATACGCTATA 1560
Qy      582 GlyThrLysValGlnLysAsnValThrAlaLysAspProGlnGlyLeuAspIleSerTyraSer 601
Db      1561 GGCATTAAGTGGCACTGTGACTGCCAGGATCCGAAAGCCTGACTGTGATTAATTC 1620
Qy      602 LeuArgGlyAspThrArgGlyTrrpleuLysIleAspHisValThrGlyGlnIlePheSer 621
Db      1621 CTGAAGACAAATAGAGAGGTTGGCTTAATATGACTCCGTCACCGGCGACATATTTAGC 1680
Qy      622 ValAlaProLeuAspArgGluAlaGlySerProTyraArgValGlnValVal 638
Db      1681 ACTGCTCCGCTGACAGGGAAACAGAAAGTGTATCGATACAAAGTGTG 1731

```

Search completed: June 22, 2003, 02:41:16
 Job time : 4969 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 20:20:34 ; Search time 699 Seconds
(without alignments)
1076.720 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345
Sequence: 1 gaattccgctcgcacacg.....cctttgcgagcgacgaattc 3345

Scoring table:
IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq.101002.*
2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	100.0	3345	20	AA18166
2	3345	100.0	3345	22	AA129511
3	3345	100.0	3345	24	ABN96791
4	3345	100.0	3345	24	ABN62444
5	3345	100.0	3345	24	ABN64424
6	3314.8	99.1	3654	24	ABN97312
7	3314.8	99.1	3654	24	ABN62484
8	3314.8	99.1	3654	24	ABN67667
9	514.2	15.4	523	24	ABN37430

10	511.4	15.3	627	21	AAA16509	Human colon cancer
11	497.2	14.9	501	24	ABN38503	Human colon tumour
12	469.4	14.0	495	24	ABN36982	Human colon tumour
13	461.6	13.8	492	24	ABN36509	Human colon tumour
14	431.4	12.9	516	24	ABN44858	CDNA encoding colo
15	410.4	12.3	568	24	ABN45154	CDNA encoding colo
16	365.6	10.9	370	24	ABN36589	Human colon tumour
17	360.4	10.8	820	22	AAH33661	Human colon cancer
18	342	10.2	342	24	ABN45746	CDNA encoding colo
19	340.4	10.2	500	24	ABN45083	CDNA encoding colo
20	331.8	9.9	335	23	AA557401	CDNA #77 encoding
21	331.8	9.9	335	23	AA558866	CDNA #1542 encoding
22	331.4	9.9	346	21	AA128480	CDNA encoding huma
23	331.4	9.9	346	22	AA128480	Colon tumour relat
24	319.8	9.6	325	22	AA129276	Human colon relat
25	284.8	8.5	591	21	AA162395	Human colon cancer
26	269.6	8.1	274	24	ABN36483	Human colon tumour
27	265.2	7.9	268	24	ABN36452	Human colon tumour
28	265	7.9	265	24	ABN27599	Stomach cancer rel
29	207.8	6.2	167343	24	ABN64403	Thyroid cancer rel
30	207.8	6.2	167343	24	ABN64403	CDNA encoding huma
31	203.8	6.1	3121	24	ABN51416	Human nervous syst
32	199.6	6.0	6618	22	ABN18101	Human nervous syst
33	199.2	6.0	10614	22	AA136948	Human musculocele
34	199	5.9	7582	22	AA164000	Human polynucleoti
35	199	5.9	7582	22	AA531678	Genomic sequence #
36	199	5.9	147419	24	ABN83574	Human CDNA differe
37	197.8	5.9	18061	24	ABN61973	Colon adenocarcino
38	197.8	5.9	18061	24	ABN62480	Colon adenocarcino
39	197.8	5.9	18061	24	ABN62481	Human CDNA sequenc
40	197.4	5.9	3417	22	AAH18467	Human musculocele
41	197.4	5.9	10612	22	AA136946	Human nervous syst
42	196.8	5.9	5161	22	ABN20105	Human prostate exp
43	195.8	5.9	530	23	ABN50892	Human nervous syst
44	195.8	5.9	6562	22	ABN19716	Human nervous syst
45	195.8	5.9	32169	23	ABN43076	Genomic sequence #

ALIGNMENTS

RESULT 1	AA18166	standard; DNA; 3345 BP.
ID	AA18166	
AC	AA18166	
DT	04-MAY-1999	(first entry)
XX		
DE	Human HPT-1 protein coding sequence.	
XX		
KW	Gastro-intestinal transport receptor; binding protein; hsi; HPT1;	
KW	D2H; hPPT1; human; GI tract receptor; sucrose-isomaltase complex;	
KW	intestinal peptide-associated transporter; hypertension; diabetes;	
KW	osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;	
KW	therapeutic agent delivery; therapy; probe; ss.	
OS	Homo sapiens.	
XX		
PN	WO9851325-A2.	
XX		
PD	19-NOV-1998.	
XX		
PF	15-MAY-1998;	98WO-US10088.
XX		
PR	15-MAY-1997;	97US-0046595.
XX		
PA	(CYTO-) CYTOGEN CORP.	
XX	(ELAN-) ELAN CORP PLC.	
XX		
PI	Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin JI;	
XX	Omahony DJ, Patterson CA, Singleton J;	
XX		

PI King GE, Wang T, Jiang Y;

XX WPI; 2001-441847/47.

PT Colon tumor associated proteins and nucleic acids useful for the
prevention, diagnosis and treatment of colonic cancer -

XX Claim 2; Page 466-467; 472pp; English.

XX The present invention describes colon tumour associated proteins (I) and
the polynucleotides (II) that encode them. (I) have cytostatic activity.
CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
CC (II) may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate colon tumour associated protein (TCAP)
CC expression, such as colonic cancer. For example, (I) and (II) may be
CC used to treat disorders associated with decreased expression by
CC rectifying mutations or deletions in a patient's genome that affect the
CC activity of TCAPs by expressing inactive proteins or to supplement the
CC patients own production of them. Additionally, (II) may be used to
CC produce the TCAP proteins, by inserting the nucleic acids into a host
CC cell culturing the cell to express the protein. (II) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC polymerase chain reaction (PCR) and hybridisation assays to detect and
CC quantitate the presence of similar nucleic acids in samples, and (I) may
CC therefore which patients may be in need of restorative therapy. (I) may
CC also be used as antigens in the production of antibodies against TCAPs
CC and in assays to identify modulators of TCAP expression and activity.
CC Anti-(I) antibodies and antagonists may also be used to down regulate
CC TCAP expression and activity. The anti-(I) antibodies may also be used
CC as diagnostic agents for detecting the presence of TCAPs in samples
CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
CC and AA24494 to AA24523 represent nucleotide and amino acid sequences
CC given in the exemplification of the present invention.

XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Query Match 100.0%; Score 3345; DB 22; Length 3345;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCGTCGACCACTGATGAGAAAGAAAGCACTTTTAAACCACTTTGTGACTTA 60
DB 1 GAATTCGCGTCGACCACTGATGAGAAAGAAAGCACTTTTAAACCACTTTGTGACTTA 60
QY 61 CAGAAAGAAATTTGAATTAAGAAAACTATGATCTTCAGGCCCATCTTCCCTGTGT 120
DB 61 CAGAAAGAAATTTGAATTAAGAAAACTATGATCTTCAGGCCCATCTTCCCTGTGT 120
QY 121 CTTCTTATGCTTTATTTGGCAACTGATATGAGCCAAAGGGGAAAGTTTAGTGACCCCTG 180
DB 121 CTTCTTATGCTTTATTTGGCAACTGATATGAGCCAAAGGGGAAAGTTTAGTGACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGAACCCAGTCAATTAATTTCCAGTTT 240
DB 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGAACCCAGTCAATTAATTTCCAGTTT 240
QY 241 AAGGCCAATCTCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACAAATATTTGTG 300
DB 241 AAGGCCAATCTCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACAAATATTTGTG 300
QY 241 AAGGCCAATCTCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACAAATATTTGTG 300
DB 241 AAGGCCAATCTCTCTGCTGTGACTTTTGAATTAAGTGGGAGACAGACAAATATTTGTG 300
QY 301 ATAGAAGGGAGGAGACTCTGTATTTACAAGAGCCTTGAGACAGGAGAAACAAGATCTACT 360
DB 301 ATAGAAGGGAGGAGACTCTGTATTTACAAGAGCCTTGAGACAGGAGAAACAAGATCTACT 360
QY 361 CACAATCTCCAGTTTGCAGCCCTGAGACGCTAATGGAATTAATAGTGAAGGTCAGTCCCT 420
DB 361 CACAATCTCCAGTTTGCAGCCCTGAGACGCTAATGGAATTAATAGTGAAGGTCAGTCCCT 420
QY 421 ATCCACATAGAAGTGAAGAGACATCAAGACATGACACCACTTTTCTCAATCAAAAGTAC 480
DB 421 ATCCACATAGAAGTGAAGAGACATCAAGACATGACACCACTTTTCTCAATCAAAAGTAC 480
QY 481 GAAGGCTCAGTAAGGACGAACCTCGCCAGGAAAGCCCTTCTGTATGTCAATGCGACA 540

DB 481 GAAGGCTCAGTAAGGACGAACCTCGCCAGGAAAGCCCTTCTGTATGTCAATGCGACA 540
QY 541 GACCTGATGATCCGGCCCACTCCCAATGGCCAGCTTTATTAACAATTTGATCCAGCTT 600
DB 541 GACCTGATGATCCGGCCCACTCCCAATGGCCAGCTTTATTAACAATTTGATCCAGCTT 600
QY 601 CCATGATCAACAATGATCATGTACTTTCAGATCAACAACAAACGGGAGCCATCTCTCT 660
DB 601 CCATGATCAACAATGATCATGTACTTTCAGATCAACAACAAACGGGAGCCATCTCTCT 660
QY 661 ACCGAGAGGATCTCAGGAATTTGAATTCCTGTAAGATCTTCTTAATCTTGATGATC 720
DB 661 ACCGAGAGGATCTCAGGAATTTGAATTCCTGTAAGATCTTCTTAATCTTGATGATC 720
QY 721 TCGATGAAGACATGAGGAGCCGAGATGAGAAATTCCTTCAGATACCAATCTGAGAT 780
DB 721 TCGATGAAGACATGAGGAGCCGAGATGAGAAATTCCTTCAGATACCAATCTGAGAT 780
QY 781 ATCATATGTGACAGAAATATTTTGAAGACCAAAACCTGTGAGATGTGAAAACTCA 840
DB 781 ATCATATGTGACAGAAATATTTTGAAGACCAAAACCTGTGAGATGTGAAAACTCA 840
QY 841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCGGTGCACATAT 900
DB 841 ACTGATCTTCACCCCATCAAAATCACTCAGGTGCGGTGGAATGATCCGGTGCACATAT 900
QY 901 TCCCTTATGTCACAAAGAAAGTCCCAATTTCCATTTTCAATTAACCAAGAAAGAT 960
DB 901 TCCCTTATGTCACAAAGAAAGTCCCAATTTCCATTTTCAATTAACCAAGAAAGAT 960
QY 961 ATTTAGTGTGACGAGCCCTTGAGACCGAAGAAAGATGATGATTTTATGACGTT 1020
DB 961 ATTTAGTGTGACGAGCCCTTGAGACCGAAGAAAGATGATGATTTTATGACGTT 1020
QY 1021 GCAAAGATGATGACGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
DB 1021 GCAAAGATGATGACGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
QY 1081 GATATTTATGATTAATCACTACATGTCGTCACAGTAAACCGTATTTGAGGTCAGGAG 1140
DB 1081 GATATTTATGATTAATCACTACATGTCGTCACAGTAAACCGTATTTGAGGTCAGGAG 1140
QY 1141 AATGAAGACATGAGTAAAGATATGAGGACCTTACTGACATGACAGGATGAAAGAAAT 1200
DB 1141 AATGAAGACATGAGTAAAGATATGAGGACCTTACTGACATGACAGGATGAAAGAAAT 1200
QY 1201 ACTGCCAAGATTTTCTTAACTACAGAGATTTGAGAGCAACCTCCAAACTTCCATGAT 1260
DB 1201 ACTGCCAAGATTTTCTTAACTACAGAGATTTGAGAGCAACCTCCAAACTTCCATGAT 1260
QY 1261 GGAATCTTCTTAACTACAGAGATTTGAGAGCAACCTCCAAACTTCCATGAT 1320
DB 1261 GGAATCTTCTTAACTACAGAGATTTGAGAGCAACCTCCAAACTTCCATGAT 1320
QY 1321 AAGCAAGATATCTCTCTGATCAAACTTACGATTAAGAGTGTCTGACAAAGATTTCAAGACC 1380
DB 1321 AAGCAAGATATCTCTCTGATCAAACTTACGATTAAGAGTGTCTGACAAAGATTTCAAGACC 1380
QY 1381 CTTTGTTTTGTGCAAACTGATTAATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
DB 1381 CTTTGTTTTGTGCAAACTGATTAATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
QY 1441 TCGATATTAAGAAACCTGACTCTTGTGAGAACCAAACTTGGGTCCACATCTTAAC 1500
DB 1441 TCGATATTAAGAAACCTGACTCTTGTGAGAACCAAACTTGGGTCCACATCTTAAC 1500
QY 1501 ATCCAGGCACTGATGCTGATGAGCCATTAAGTGGAGTCTTAATATTTCTGATCATATC 1560
DB 1501 ATCCAGGCACTGATGCTGATGAGCCATTAAGTGGAGTCTTAATATTTCTGATCATATC 1560
QY 1561 ATAAAGGAGACATGAGGAGCCCTGAGGAGTTTGAACAGATCCCATACCAACCGGA 1620

Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTGACACAGATCCCATACCAACACCGGA 1620
 Qy 1621 TATGTCATTAATTAATAAAGCCTCTGATTTTGAACAGACAGCTGTTCCAACTGTGTTC 1680
 Db 1621 TATGTCATTAATTAATAAAGCCTCTGATTTTGAACAGACAGCTGTTCCAACTGTGTTC 1680
 Qy 1681 AAAGCAGAAAATCTGAGCCTCTAGTGTGTTGGTGAAGTCAATGCAAGTCTTTTGGC 1740
 Db 1681 AAAGCAGAAAATCTGAGCCTCTAGTGTGTTGGTGAAGTCAATGCAAGTCTTTTGGC 1740
 Qy 1741 AAGTTCACGCTTAATTTGACAGATGTGAATGACACCTCAATTTTCCCAACAGCTATTC 1800
 Db 1741 AAGTTCACGCTTAATTTGACAGATGTGAATGACACCTCAATTTTCCCAACAGCTATTC 1800
 Qy 1801 CAAGCAGAAAGTCAGTAGAGATGTAGCTAATAGGCACTAAAGTGGGCAATGTGACTGCCAAG 1860
 Db 1801 CAAGCAGAAAGTCAGTAGAGATGTAGCTAATAGGCACTAAAGTGGGCAATGTGACTGCCAAG 1860
 Qy 1861 GATCCAGAAAGTCTGACATTAAGCTATTCACTGAGGGAGACACAGAGGTTGGCTTAAA 1920
 Db 1861 GATCCAGAAAGTCTGACATTAAGCTATTCACTGAGGGAGACACAGAGGTTGGCTTAAA 1920
 Qy 1921 ATTGACACGCTGATGCTGTGAGATCTTTAGTGTGCTCATTTGACAGAGAACCGGAAGT 1980
 Db 1921 ATTGACACGCTGATGCTGTGAGATCTTTAGTGTGCTCATTTGACAGAGAACCGGAAGT 1980
 Qy 1981 CCATATCGGGGTACAAAGTGTGTGGCCACAGAAAGTAGGGGGGTCTTCTTAAGCTGTGTCA 2040
 Db 1981 CCATATGGGGTACAAAGTGTGTGGCCACAGAAAGTAGGGGGGTCTTCTTAAGCTGTGTCA 2040
 Qy 2041 GAGTTCACCTGATCTTATGATGTGAATGACAAACCTCCAGGCTAGCCAAAGACTAC 2100
 Db 2041 GAGTTCACCTGATCTTATGATGTGAATGACAAACCTCCAGGCTAGCCAAAGACTAC 2100
 Qy 2101 ACGGGCTTGTCTTCTTCCCATCCCTCAGTGCACCTGAAAGTCTCATTTTCCAGGCTACT 2160
 Db 2101 ACGGGCTTGTCTTCTTCCCATCCCTCAGTGCACCTGAAAGTCTCATTTTCCAGGCTACT 2160
 Qy 2161 GATGATGATGACGACTTAATTTCCGGGGTCCCATTTTCAATTTTCCCTGGGAGTGGAAAGC 2220
 Db 2161 GATGATGATGACGACTTAATTTCCGGGGTCCCATTTTCAATTTTCCCTGGGAGTGGAAAGC 2220
 Qy 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCAATGTGACTCATGCCGACGCTCTACGAG 2280
 Db 2221 TTACAAAACGACTGGGAAGTTTCCAAAATCAATGTGACTCATGCCGACGCTCTACGAG 2280
 Qy 2281 CACACAGACTTTGAGAGAGAGGGCGTATGTCTTATCCGATCAATGATGGGGGTGCG 2340
 Db 2281 CACACAGACTTTGAGAGAGAGGGCGTATGTCTTATCCGATCAATGATGGGGGTGCG 2340
 Qy 2341 CCAACCTTGAAGGCACTGTTCTTACAGTTACATTTGCAAGTTGTGTGGAAGAGT 2400
 Db 2341 CCAACCTTGAAGGCACTGTTCTTACAGTTACATTTGCAAGTTGTGTGGAAGAGT 2400
 Qy 2401 TGTTCGGGCGACAGGTCAACAGACTGGATACCACTGTGGGCAATGGAGTTGGTATA 2460
 Db 2401 TGTTCGGGCGACAGGTCAACAGACTGGATACCACTGTGGGCAATGGAGTTGGTATA 2460
 Qy 2461 CTGCTGACCAACCTTCTGTGTGATGTGTATATTTAGCAGTTGTGTATCCGATTAAG 2520
 Db 2461 CTGCTGACCAACCTTCTGTGTGATGTGTATATTTAGCAGTTGTGTATCCGATTAAG 2520
 Qy 2521 AAGGATTAAGGCAAGATTAATGTGAAGTCAAGCACTGAAGTCAAACTCTGAGA 2580
 Db 2521 AAGGATTAAGGCAAGATTAATGTGAAGTCAAGCACTGAAGTCAAACTCTGAGA 2580
 Qy 2581 AGCTGAATTTGAAAGGAATGTTGAATTTATATAGCAAGTGTATTTAGCAACAACA 2640
 Db 2581 AGCTGAATTTGAAAGGAATGTTGAATTTATATAGCAAGTGTATTTAGCAACAACA 2640
 Qy 2641 TCTCATCTCTATTACTTTTCACTACAGTGCATTAATTTTAAACAGATATTCCTCT 2700
 Db 2641 TCTCATCTCTATTACTTTTCACTACAGTGCATTAATTTTAAACAGATATTCCTCT 2700

Qy 2701 TGTCTTAAATATTGCTAAATATTCTTTTGAAGTGAAGTCTGTGCGCCAG 2760
 Db 2701 TGTCTTAAATATTGCTAAATATTCTTTTGAAGTGAAGTCTGTGCGCCAG 2760
 Qy 2761 GCTGAGATACAGTGTGTGATCCAGCTCACTCAACCTCCGCTCTGGGTTCAATGA 2820
 Db 2761 GCTGAGATACAGTGTGTGATCCAGCTCACTCAACCTCCGCTCTGGGTTCAATGA 2820
 Qy 2821 TTCTCCGCTCAGCTTCCTAAGTCTGGGTTTACAGGACCCACCAATGCCCCAGCT 2880
 Db 2821 TTCTCCGCTCAGCTTCCTAAGTCTGGGTTTACAGGACCCACCAATGCCCCAGCT 2880
 Qy 2881 AATTTTGTATTTTAAATAGACAGGGGTTTGGCAATTTGGCAGGCTGTGCTTGAAGT 2940
 Db 2881 AATTTTGTATTTTAAATAGACAGGGGTTTGGCAATTTGGCAGGCTGTGCTTGAAGT 2940
 Qy 2941 CTGACGTCAGATGATCTGCTGCTGCTGTGCTCCCAATACAGGCATGAACCACTGAC 3000
 Db 2941 CTGACGTCAGATGATCTGCTGCTGCTGTGCTCCCAATACAGGCATGAACCACTGAC 3000
 Qy 3001 CCTACTAGATATTTCATGTGCTATAGCACTTGAAGATTTTTCATTTTCCATGACAT 3060
 Db 3001 CCTACTAGATATTTCATGTGCTATAGCACTTGAAGATTTTTCATTTTCCATGACAT 3060
 Qy 3061 TTTTCTCTCTGCAAAATGCGTACTGCTGCTGCTTCCCTTTGGGGCAAGACACT 3120
 Db 3061 TTTTCTCTCTGCAAAATGCGTACTGCTGCTGCTTCCCTTTGGGGCAAGACACT 3120
 Qy 3121 CATTAATAATCTGTACATTTTCTTTATCAAGAGATATATCAGTGTGTCTCATAGA 3180
 Db 3121 CATTAATAATCTGTACATTTTCTTTATCAAGAGATATATCAGTGTGTCTCATAGA 3180
 Qy 3181 ACTGCTGATTCATTTATGTTTTCGATTCATCTGTGTCCCTTCACTCTTGTAC 3240
 Db 3181 ACTGCTGATTCATTTATGTTTTCGATTCATCTGTGTCCCTTCACTCTTGTAC 3240
 Qy 3241 TCCTTGGTATTTACAGTGAATTTCAAACTTTGTCAGAGAGAAAAGTGAAGTCTGAG 3300
 Db 3241 TCCTTGGTATTTACAGTGAATTTCAAACTTTGTCAGAGAGAAAAGTGAAGTCTGAG 3300
 Qy 3301 GAAAAATTAATAATAAAGAACAGCCTTTTGGGCGCGGAATTC 3345
 Db 3301 GAAAAATTAATAATAAAGAACAGCCTTTTGGGCGCGGAATTC 3345

RESULT 3
 AEN96791
 ID AEN96791 standard; DNA; 3345 BP.
 XX
 AC AEN96791;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Gene #13289 used to diagnose liver cancer.
 XX
 DE Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
 KW metastatic liver tumour; cytostatic; expression profile; disease state;
 KW disease progression; drug toxicity; drug efficacy; drug metabolism.
 OS Homo sapiens.
 XX
 PN W0200229103-A2.
 XX
 PD 11-APR-2002.
 XX
 PF 02-OCT-2001; 2001WC-US30589.
 XX
 PR 02-OCT-2000; 2000US-237054P.
 XX
 PA (GENE-) GENE LOGIC INC.
 XX
 PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX
DR WPI; 2002-426119/45.
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
PS Claim 1; SEQ ID NO 3289; 298bp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIP0
CC at ftp.wipo.int/pub/published_pct_sequences..
XX
SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
Query Match 100.0%; Score 3345; DB 24; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATTCGCTCGACCACTGATGGAAGAAAGACTTTTAAACCACTTTTGGACTTA 60
DB 1 GAATTCGCTCGACCACTGATGGAAGAAAGACTTTTAAACCACTTTTGGACTTA 60
QY 61 CAGAAAGAAATTTGAATTAAGAAAGAACTATGATACCTTCAGGCCCATCTTCCTGCTGT 120
DB 61 CAGAAAGAAATTTGAATTAAGAAAGAACTATGATACCTTCAGGCCCATCTTCCTGCTGT 120
QY 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGGAAAGTTTATGAGACCCCTG 180
DB 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGGAAAGTTTATGAGACCCCTG 180
QY 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGGAAAGTTTATGAGACCCCTG 180
DB 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGGAAAGTTTATGAGACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTTT 240
DB 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTTT 240
QY 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTTT 240
DB 181 AAACCCATGACATTTTCTATTTATGAAGCCCAAGACCGAGTCAATTTATTTCCAGTTT 240
QY 241 AAGGCCAATCTCTGCTGCTGACTTTTGAATTAAGTGGGAGAGACAATATTTGTG 300
DB 241 AAGGCCAATCTCTGCTGCTGCTGACTTTTGAATTAAGTGGGAGAGACAATATTTGTG 300
QY 241 AAGGCCAATCTCTGCTGCTGCTGACTTTTGAATTAAGTGGGAGAGACAATATTTGTG 300
DB 241 AAGGCCAATCTCTGCTGCTGCTGACTTTTGAATTAAGTGGGAGAGACAATATTTGTG 300
QY 301 ATAGAACGGAGGAGACTTCTGTATTTACAACAGAGCTTTGACAGGGAAACAAGATCTACT 360
DB 301 ATAGAACGGAGGAGACTTCTGTATTTACAACAGAGCTTTGACAGGGAAACAAGATCTACT 360
QY 301 ATAGAACGGAGGAGACTTCTGTATTTACAACAGAGCTTTGACAGGGAAACAAGATCTACT 360
DB 301 ATAGAACGGAGGAGACTTCTGTATTTACAACAGAGCTTTGACAGGGAAACAAGATCTACT 360
QY 361 CACAAATCTCCAGGTTTGACGCTTGAAGCTTATGGAATTTATGAGAGGCTCCAGTCCCT 420
DB 361 CACAAATCTCCAGGTTTGACGCTTGAAGCTTATGGAATTTATGAGAGGCTCCAGTCCCT 420
QY 361 CACAAATCTCCAGGTTTGACGCTTGAAGCTTATGGAATTTATGAGAGGCTCCAGTCCCT 420
DB 361 CACAAATCTCCAGGTTTGACGCTTGAAGCTTATGGAATTTATGAGAGGCTCCAGTCCCT 420
QY 421 ATGACCAATAGAAAGTGAAGAGCATCAACAGCAATGCAAGCTTTCTCAGTCAAAAGTAC 480
DB 421 ATGACCAATAGAAAGTGAAGAGCATCAACAGCAATGCAAGCTTTCTCAGTCAAAAGTAC 480
QY 421 ATGACCAATAGAAAGTGAAGAGCATCAACAGCAATGCAAGCTTTCTCAGTCAAAAGTAC 480
DB 421 ATGACCAATAGAAAGTGAAGAGCATCAACAGCAATGCAAGCTTTCTCAGTCAAAAGTAC 480
QY 481 GAAAGCTCAGTAAAGCAGAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCATGACCA 540
DB 481 GAAAGCTCAGTAAAGCAGAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCATGACCA 540
QY 481 GAAAGCTCAGTAAAGCAGAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCATGACCA 540
DB 481 GAAAGCTCAGTAAAGCAGAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCATGACCA 540
QY 541 GACCTGGATATCCGGCCCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
DB 541 GACCTGGATATCCGGCCCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
QY 541 GACCTGGATATCCGGCCCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
DB 541 GACCTGGATATCCGGCCCACTCCCAATGCGCAGCTTTATTAACAGATTTGATCCAGTT 600
QY 601 CCATATGATCAACAAATGTATGATTTGATGATCAACAAACAAAGGAGCATCTCTCTT 660

DB 601 CCATATGATCAACAAATGTATGATTTGATGATCAACAAACAAAGGAGCATCTCTCTT 660
QY 661 ACCCGAGAGGAGATCTCAGGAATTTGATCTCTGCTTAAGAAATCTTCTATATCTGGATC 720
DB 661 ACCCGAGAGGAGATCTCAGGAATTTGATCTCTGCTTAAGAAATCTTCTATATCTGGATC 720
QY 721 TCAGTGAAGACATGAGGAGCCAGAGTGAATTTCTTCAGTATACCAATCTGTGAT 780
DB 721 TCAGTGAAGACATGAGGAGCCAGAGTGAATTTCTTCAGTATACCAATCTGTGAT 780
QY 781 ATCATATGAGCAGAGAAATTTTGAAGAGCAACAAACCTGTGAGATGTTGAAACTCA 840
DB 781 ATCATATGAGCAGAGAAATTTTGAAGAGCAACAAACCTGTGAGATGTTGAAACTCA 840
QY 841 ACTGATCTTCAACCCCATCAAAATCACTCAGGTGGGTGGAATGATCCGGTGACAAAT 900
DB 841 ACTGATCTTCAACCCCATCAAAATCACTCAGGTGGGTGGAATGATCCGGTGACAAAT 900
QY 901 TCCTTAGTTGACAAAGAGAGTGCAGAGATTCCTCAATTTGATGACAGAGAGAT 960
DB 901 TCCTTAGTTGACAAAGAGAGTGCAGAGATTCCTCAATTTGATGACAGAGAGAT 960
QY 961 ATTTAGTGAATGAGCCCTTGGACCGAGAGAAAGATGATATGTTTATGAGATT 1020
DB 961 ATTTAGTGAATGAGCCCTTGGACCGAGAGAAAGATGATATGTTTATGAGATT 1020
QY 1021 GCAAGAGATGATGAGGAGAAACCACTTTGATATTCCTGGAATTCATGTAAGTTAA 1080
DB 1021 GCAAGAGATGATGAGGAGAAACCACTTTGATATTCCTGGAATTCATGTAAGTTAA 1080
QY 1081 GATATTATGATATATCACTATGATGCTGACAGTAACTGATTTAGAGTCCAGAG 1140
DB 1081 GATATTATGATATATCACTATGATGCTGACAGTAACTGATTTAGAGTCCAGAG 1140
QY 1141 AATGAAGACTGGGTAACAGTATGCGGACCTTACTGACATGACAGGATGAAGAAAT 1200
DB 1141 AATGAAGACTGGGTAACAGTATGCGGACCTTACTGACATGACAGGATGAAGAAAT 1200
QY 1201 ACTGCCAAGCTTTTCTAACTCAAGATTTGAGAGCAAACTCCCAATCCAGAT 1260
DB 1201 ACTGCCAAGCTTTTCTAACTCAAGATTTGAGAGCAAACTCCCAATCCAGAT 1260
QY 1261 GGAATCTTCTAATCCAACTATGCTGAAATGTTACGTTAGCTAAACGTCCTTGAAG 1320
DB 1261 GGAATCTTCTAATCCAACTATGCTGAAATGTTACGTTAGCTAAACGTCCTTGAAG 1320
QY 1321 AAGCAAGATCTCTCAGTACAACTTAAACGATAGAGGTGCTGACAAAGATTTCAAGACC 1380
DB 1321 AAGCAAGATCTCTCAGTACAACTTAAACGATAGAGGTGCTGACAAAGATTTCAAGACC 1380
QY 1381 CTTGTTTCTGGAATCAACGTTATGATATCAATGATATCCCATCTTGAAGAA 1440
DB 1381 CTTGTTTCTGGAATCAACGTTATGATATCAATGATATCCCATCTTGAAGAA 1440
QY 1441 TCAGATTAAGAAACCTGACTCTTGTGAGAGACAAACCTTGGTCCACATCTTAAACC 1500
DB 1441 TCAGATTAAGAAACCTGACTCTTGTGAGAGACAAACCTTGGTCCACATCTTAAACC 1500
QY 1501 ATCCAGGCACTGATGCTGATGAGCAATTTACGAGGAGTTCTTAAATTCGTATCATATC 1560
DB 1501 ATCCAGGCACTGATGCTGATGAGCAATTTACGAGGAGTTCTTAAATTCGTATCATATC 1560
QY 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
DB 1561 ATTAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
QY 1621 TATGTCATTAATTAAGAGCTCTTGAATTTGAAACAGCAGCTGTTTCCAACTATGTTCC 1680
DB 1621 TATGTCATTAATTAAGAGCTCTTGAATTTGAAACAGCAGCTGTTTCCAACTATGTTCC 1680
QY 1681 AAAGCAGAAATCTGAGCTCTGATGTTTGGTGAAGTACAATGCAAGTCTTTTGGC 1740

Db 1681 AAAGAGAAAACTGACCTCTAGTGTGGTGTGAAGTACAAATGCAAGTCTTTTGGC 1740
 Qy 1741 AAGTTCAAGCTTATTTGACAGATGTGAATGAGACCTCAATTTTCCCAACAGTATTC 1800
 Db 1741 AAGTTCAAGCTTATTTGACAGATGTGAATGAGACCTCAATTTTCCCAACAGTATTC 1800
 Qy 1801 CAAGCGAAGTCAAGTGAAGATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
 Db 1801 CAAGCGAAGTCAAGTGAAGATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
 Qy 1861 GATCAGAGGCTGAGCATATAGCTATTCATGAGGGGAGACAGAGAGTGTGCTTAA 1920
 Db 1861 GATCAGAGGCTGAGCATATAGCTATTCATGAGGGGAGACAGAGAGTGTGCTTAA 1920
 Qy 1921 ATTGACCAAGTCAAGTGAAGATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
 Db 1921 ATTGACCAAGTCAAGTGAAGATGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1980
 Qy 1981 CCAATATGCGGATCAAGTGTGGGACAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040
 Db 1981 CCAATATGCGGATCAAGTGTGGGACAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2040
 Qy 2041 GAGTTCCAGCTGATCCCTTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2100
 Db 2041 GAGTTCCAGCTGATCCCTTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2100
 Qy 2101 ACGGCTTGTCTTCTGCGCATCCCTCAGTCACTGGAAGTCAATTTTGAAGTCACT 2160
 Db 2101 ACGGCTTGTCTTCTGCGCATCCCTCAGTCACTGGAAGTCAATTTTGAAGTCACT 2160
 Qy 2161 GATGATGATGACAGCTTATTTGCGGAGTCCCATTTTCAATTTTCCCTGCGAGTGAAGC 2220
 Db 2161 GATGATGATGACAGCTTATTTGCGGAGTCCCATTTTCAATTTTCCCTGCGAGTGAAGC 2220
 Qy 2221 TTACAAAACAGTCTGGAAGTTCGAAATCAATGATCAATGCTCAATGCTCAATGCTCAATG 2280
 Db 2221 TTACAAAACAGTCTGGAAGTTCGAAATCAATGATCAATGCTCAATGCTCAATGCTCAATG 2280
 Qy 2281 CACACAGACTTTGAGAGAGGAGGAGTGTGCTTGAATCCGATCAATGATGAGGAGTCCG 2340
 Db 2281 CACACAGACTTTGAGAGAGGAGGAGTGTGCTTGAATCCGATCAATGATGAGGAGTCCG 2340
 Qy 2341 CCACCTTTGAGAGGAGTGTGCTTGAATCCGATCAATGATGAGGAGTCCG 2400
 Db 2341 CCACCTTTGAGAGGAGTGTGCTTGAATCCGATCAATGATGAGGAGTCCG 2400
 Qy 2401 TGTTCGCGGACAGAGTCAACAGCTGCGATCCCACTGTGGGACATGGGAGTGTGATTA 2460
 Db 2401 TGTTCGCGGACAGAGTCAACAGCTGCGATCCCACTGTGGGACATGGGAGTGTGATTA 2460
 Qy 2461 CTGCTGACCAACCTTCTGAGTGTGATTAATTTTACAGTGTGATTAATTTTACAGTGTGATTA 2520
 Db 2461 CTGCTGACCAACCTTCTGAGTGTGATTAATTTTACAGTGTGATTAATTTTACAGTGTGATTA 2520
 Qy 2521 AAGGATTAAGCAAGATTAATTTTGAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2580
 Db 2521 AAGGATTAAGCAAGATTAATTTTGAAGTCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2580
 Qy 2581 AGCTGAATTTGAAGAGATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
 Db 2581 AGCTGAATTTGAAGAGATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2640
 Qy 2641 TGTTCATCTATTTCTTTCATCTTCAAGTCAATTAATTTTGAAGTCAATTAATTTTGAAGTCA 2700
 Db 2641 TGTTCATCTATTTCTTTCATCTTCAAGTCAATTAATTTTGAAGTCAATTAATTTTGAAGTCA 2700
 Qy 2701 TGTTCATCTATTTCTTTCATCTTCAAGTCAATTAATTTTGAAGTCAATTAATTTTGAAGTCA 2760
 Db 2701 TGTTCATCTATTTCTTTCATCTTCAAGTCAATTAATTTTGAAGTCAATTAATTTTGAAGTCA 2760
 Qy 2761 GCTGAGATCAAGTGTGATTCAGTCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2820
 Db 2761 GCTGAGATCAAGTGTGATTCAGTCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 2820

Qy 2821 TTCTCTGCTCAGCTTCTTAAGTGTGGGTTTACAGGACCCACCACTGCCCAGCT 2880
 Db 2821 TTCTCTGCTCAGCTTCTTAAGTGTGGGTTTACAGGACCCACCACTGCCCAGCT 2880
 Qy 2881 AATTTTGTATTTTAAATTAAGAGAGGAGTGTGGGTTTACAGGAGTGTGGGTTTACAGG 2940
 Db 2881 AATTTTGTATTTTAAATTAAGAGAGGAGTGTGGGTTTACAGGAGTGTGGGTTTACAGG 2940
 Qy 2941 CTGACGTCAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Db 2941 CTGACGTCAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3000
 Qy 3001 CCTACTAGATATTTTCAATGCTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3060
 Db 3001 CCTACTAGATATTTTCAATGCTATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3060
 Qy 3061 TTTTCTCTCTGCAAAATGAGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
 Db 3061 TTTTCTCTCTGCAAAATGAGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
 Qy 3121 CATTAAATATTTCTGATATTTTCTTTATCAAGAGATTAATCAAGTGTGTGTGTGTGTGTGT 3180
 Db 3121 CATTAAATATTTCTGATATTTTCTTTATCAAGAGATTAATCAAGTGTGTGTGTGTGTGTGT 3180
 Qy 3181 ACTGCGTGAATTCATTTATGTTTCTGATTCATTCGATTCGATTCGATTCGATTCGATTCGAT 3240
 Db 3181 ACTGCGTGAATTCATTTATGTTTCTGATTCATTCGATTCGATTCGATTCGATTCGATTCGAT 3240
 Qy 3241 TCCCTTGTGATTTTCACTGAATTTTCAACATTTTCAAGAGATTAATCAAGTGTGTGTGTGT 3300
 Db 3241 TCCCTTGTGATTTTCACTGAATTTTCAACATTTTCAAGAGATTAATCAAGTGTGTGTGTGT 3300
 Qy 3301 GAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3345
 Db 3301 GAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3345

RESULT 4
 ABL62444
 ID ABL62444 standard; DNA; 3345 BP.
 XX
 AC ABL62444;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Colon adenocarcinoma related gene sequence SEQ ID NO: 781.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antileukemia; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN NC0200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001MO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.

PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX (AVALON PHARM.
 PA
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
 PI Sopet DR, Weaver Z;
 XX WPI: 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 781; 44pp; English.
 PS
 XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (1) of a signature gene set, where (1)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL6164
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (1) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX
 SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
 Query Match 100.0%; Score 3345; DB 24; Length 3345;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAATTCGCTCGACCACTGAATGGAAGAAAGACCTTTAAACCACTTTTGACCTTA 60
 DB 1 GAATTCGCTCGACCACTGAATGGAAGAAAGACCTTTAAACCACTTTTGACCTTA 60
 QY 61 CAGAAAGAAATTGGAATAAGAAAACCTATGATCTTACGGCCCACTTCCTCCTGTGT 120

DB 61 CAGAAAGAAATTGGAATAAGAAAACCTATGATCTTACGGCCCACTTCCTCCTGTGT 120
 QY 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAGTTTGTGACCCCTG 180
 DB 121 CTTCTTATGCTTTATTTGGCACTGATATGCGCAAGAGGGAGTTTGTGACCCCTG 180
 QY 181 AATCCATGACATTTTCTATTTATGAAAGGCAAGAACCGAGTCAATTAATATCCAGTTT 240
 DB 181 AATCCATGACATTTTCTATTTATGAAAGGCAAGAACCGAGTCAATTAATATCCAGTTT 240
 QY 241 AAGGCCAATCTCTCTGCTGATGACTTTTGAATTAATGCGGAGACAGCAATATTTGTG 300
 DB 241 AAGGCCAATCTCTCTGCTGATGACTTTTGAATTAATGCGGAGACAGCAATATTTGTG 300
 QY 301 ATGAAACGGAGGAGACTTCTGTATTAACAAGAGCCTTGGACAGGAAACAAATCTACT 360
 DB 301 ATGAAACGGAGGAGACTTCTGTATTAACAAGAGCCTTGGACAGGAAACAAATCTACT 360
 QY 361 CACAATCTCCAGGTTGCAAGCCCTGAGCGCTAATGGAATTAATGAGAGGTCAGTCCCT 420
 DB 361 CACAATCTCCAGGTTGCAAGCCCTGAGCGCTAATGGAATTAATGAGAGGTCAGTCCCT 420
 QY 421 ATCACCATAAGATGAAGACATCAACGACATGACCCAGTTTCTCAGTCAAGTAC 480
 DB 421 ATCACCATAAGATGAAGACATCAACGACATGACCCAGTTTCTCAGTCAAGTAC 480
 QY 481 GAAGGCTCAGTAAAGGAGAACTCTGCGCCGAGAAACCCCTTCTTATGTCATGTCACA 540
 DB 481 GAAGGCTCAGTAAAGGAGAACTCTGCGCCGAGAAACCCCTTCTTATGTCATGTCACA 540
 QY 541 GACCTGATATATCGGCGCACTCCCAATGCGCAGCTTTATACAGATTTGTCATCAGCTT 600
 DB 541 GACCTGATATATCGGCGCACTCCCAATGCGCAGCTTTATACAGATTTGTCATCAGCTT 600
 QY 601 CCCATGATCAACATGTCATGATCTTTCAGATCAACAAACGGAGCCATCTCTCT 660
 DB 601 CCCATGATCAACATGTCATGATCTTTCAGATCAACAAACGGAGCCATCTCTCT 660
 QY 661 ACCCGAGAGGATCTCAGAAATTTGAATTCCTGCTTAAGATCCCTTATCTGAGATC 720
 DB 661 ACCCGAGAGGATCTCAGAAATTTGAATTCCTGCTTAAGATCCCTTATCTGAGATC 720
 QY 721 TCAATGAGGACATGAGGAGCCAGAGTGAATTCCTTCAGTATACCATCTGTGAT 780
 DB 721 TCAATGAGGACATGAGGAGCCAGAGTGAATTCCTTCAGTATACCATCTGTGAT 780
 QY 781 ATCATAGTGAACAGAAATATTTGAAAGCAACAAACCTGTGAGATGTGAAACTCA 840
 DB 781 ATCATAGTGAACAGAAATATTTGAAAGCAACAAACCTGTGAGATGTGAAACTCA 840
 QY 841 ACTGATCTTCAACCCATCAAAATCACTCAGGTGCGGTGAATATCCCGGTGCAATAT 900
 DB 841 ACTGATCTTCAACCCATCAAAATCACTCAGGTGCGGTGAATATCCCGGTGCAATAT 900
 QY 901 TCCCTAGTTGCAAGAAGAGTGCAGAAATTCCTTCAATTTGACAGGAGAT 960
 DB 901 TCCCTAGTTGCAAGAAGAGTGCAGAAATTCCTTCAATTTGACAGGAGAT 960
 QY 961 ATTTACGTGACTGAGCCCTTGACCCGAGAAAGAAAGATGATGTTTTTATGACGTT 1020
 DB 961 ATTTACGTGACTGAGCCCTTGACCCGAGAAAGAAAGATGATGTTTTTATGACGTT 1020
 QY 1021 GCAAAAGATGATGAGGAAACCACTTTCAATTCCTGGAATTCATGTAAGTAAAGTAA 1080
 DB 1021 GCAAAAGATGATGAGGAAACCACTTTCAATTCCTGGAATTCATGTAAGTAAAGTAA 1080
 QY 1081 GATATTAATGATATCAACATCAATGTCGACACAGTAACCGTATTTAGAGTCCAGAG 1140
 DB 1081 GATATTAATGATATCAACATCAATGTCGACACAGTAACCGTATTTAGAGTCCAGAG 1140
 QY 1141 AATGAGACTGAGTAAACAGATGCGGACCTTATCTGACATGACAGGATGAAGAAAT 1200
 DB 1141 AATGAGACTGAGTAAACAGATGCGGACCTTATCTGACATGACAGGATGAAGAAAT 1200

D 1141 AATGAGACTGGGTAACGATATCGGACCTTACTGCACTGACAGGGATGAAGAAAT 1200
Q 1201 ACTGCACACAGTTTCTTAAACTACAGAGATTGTGGAGAACTCCCAACTTCCCATGGAT 1260
D 1201 ACTGCCAACAGTTTCTTAAACTACAGAGATTGTGGAGAACTCCCAACTTCCCATGGAT 1260
Q 1261 GGAATCTTCTTAATCCAACTTATGCTGGAAATGTTAACAATTAAAGCTCTTGAG 1320
D 1261 GGAATCTTCTTAATCCAACTTATGCTGGAAATGTTAACAATTAAAGCTCTTGAG 1320
Q 1321 AAGCAATATCTCTCGATCACTTAACATTAAGATTAAGGTCTGACAAAGATTTCAAAACC 1380
D 1321 AAGCAATATCTCTCGATCACTTAACATTAAGATTAAGGTCTGACAAAGATTTCAAAACC 1380
Q 1381 CTTTGTGTTGTCGAATCAAGTATGATATCATGATCAGATCCCATCTTTGAAGAA 1440
D 1381 CTTTGTGTTGTCGAATCAAGTATGATATCATGATCAGATCCCATCTTTGAAGAA 1440
Q 1441 TCGATATATGAAAACCTGACTTTGCTGAAGACAAACATTTGGGTCACCATCTTAACC 1500
D 1441 TCGATATATGAAAACCTGACTTTGCTGAAGACAAACATTTGGGTCACCATCTTAACC 1500
Q 1501 ATCCAGGCCACTGATGCTGATGAGCCATTTATGAGGAGTTCTAAATTCGTATCATATC 1560
D 1501 ATCCAGGCCACTGATGCTGATGAGCCATTTATGAGGAGTTCTAAATTCGTATCATATC 1560
Q 1561 ATAAAGGAGACAGTGAAGGACGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
D 1561 ATAAAGGAGACAGTGAAGGACGCTGGGGGTTGACACAGATCCCATACCAACCGGA 1620
Q 1621 TATGTCATTAATTAAGGCTCTTGAATTTTGAACAGACGCTGTTTCAACATTTGTGTC 1680
D 1621 TATGTCATTAATTAAGGCTCTTGAATTTTGAACAGACGCTGTTTCAACATTTGTGTC 1680
Q 1681 AAAGCAGAAATCTGAGCTCTAGTCTTGGTGTGAGTGAAGTGAATGCAAGTCTTTGCC 1740
D 1681 AAAGCAGAAATCTGAGCTCTAGTCTTGGTGTGAGTGAAGTGAATGCAAGTCTTTGCC 1740
Q 1741 AAGTTCACGCTTATTTGACAGATGTGATGAAGACCTCAATTTTCCCAACAGTATTC 1800
D 1741 AAGTTCACGCTTATTTGACAGATGTGATGAAGACCTCAATTTTCCCAACAGTATTC 1800
Q 1801 CAAGCAGAAAGTCAAGTGAAGATGTAGCTAAGGACCTAAGTGGGCAATGTACTGCCAG 1860
D 1801 CAAGCAGAAAGTCAAGTGAAGATGTAGCTAAGGACCTAAGTGGGCAATGTACTGCCAG 1860
Q 1861 GATCCAGAAAGTCTGACATTAAGCTATTCCTGAAGGAGACACAAAGAGTTGGCTTAA 1920
D 1861 GATCCAGAAAGTCTGACATTAAGCTATTCCTGAAGGAGACACAAAGAGTTGGCTTAA 1920
Q 1921 ATTGACCAAGTGAAGTGAAGATCTTATGCTGGCTCAATTGACAGAGAACCGGAAGT 1980
D 1921 ATTGACCAAGTGAAGTGAAGATCTTATGCTGGCTCAATTGACAGAGAACCGGAAGT 1980
Q 1981 CCATATCGGGTACAAAGTGTGGCCACAAAGATGAGGGGCTCTCTTAAGCTCTGTGCA 2040
D 1981 CCATATCGGGTACAAAGTGTGGCCACAAAGATGAGGGGCTCTCTTAAGCTCTGTGCA 2040
Q 2041 GACTTTCACCTGATCTTATGATGTGAATGAACAACCTCCAGGCTAGCAAGAGATAC 2100
D 2041 GACTTTCACCTGATCTTATGATGTGAATGAACAACCTCCAGGCTAGCAAGAGATAC 2100
Q 2101 ACGGCTTGTCTTCTGCGCATCCCTCAGTGAAGCTGGAAGTCAATTTTGAGGGTACT 2160
D 2101 ACGGCTTGTCTTCTGCGCATCCCTCAGTGAAGCTGGAAGTCAATTTTGAGGGTACT 2160
Q 2161 GATGATGATCAGACCTTATTTGGGGGTCCTCATTTTACATTTTCCCTCGGCAAGTGAAG 2220
D 2161 GATGATGATCAGACCTTATTTGGGGGTCCTCATTTTACATTTTCCCTCGGCAAGTGAAG 2220
Q 2221 TTACAAAAGACGCTGGGAAGTTTCCAAATCAATGTATCTCATGCTCCGAGCTGTACAGG 2280
D 2221 TTACAAAAGACGCTGGGAAGTTTCCAAATCAATGTATCTCATGCTCCGAGCTGTACAGG 2280

Q 2281 CACACAGACTTTGAGAGAGAGGGGCTATGCTGCTTGAATCCGATCATGATGAGGGGTCG 2340
D 2281 CACACAGACTTTGAGAGAGAGGGGCTATGCTGCTTGAATCCGATCATGATGAGGGGTCG 2340
Q 2341 CCAACCTTGGAGAGCATTTGTTCTTTACAGATTACATTCGACGTTGTGTGAAGAGT 2400
D 2341 CCAACCTTGGAGAGCATTTGTTCTTTACAGATTACATTCGACGTTGTGTGAAGAGT 2400
Q 2401 TGTTCGCGCAGACAGTCAACGACCTGGATACCACTGTGGCATGAGAGTTGGTATA 2460
D 2401 TGTTCGCGCAGACAGTCAACGACCTGGATACCACTGTGGCATGAGAGTTGGTATA 2460
Q 2461 CTGCTGACACCTCTCGTGATTTGGTATTAATTTAGCAGTTGTGTTATCCGATTAAG 2520
D 2461 CTGCTGACACCTCTCGTGATTTGGTATTAATTTAGCAGTTGTGTTATCCGATTAAG 2520
Q 2521 AAGGATTAAGGCAAGATTAATGTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
D 2521 AAGGATTAAGGCAAGATTAATGTTGAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Q 2581 AGCTGAATTTGAAAAGGAATGTTGAATTTATATAGCAAGTCTATTTACGACAACCA 2640
D 2581 AGCTGAATTTGAAAAGGAATGTTGAATTTATATAGCAAGTCTATTTACGACAACCA 2640
Q 2641 TCTCATCTTATCTTTCTCTACGATGATTAATTTTAAACAGATATTCCTCT 2700
D 2641 TCTCATCTTATCTTTCTCTACGATGATTAATTTTAAACAGATATTCCTCT 2700
Q 2701 TGTCTTATTAATTTTCTTAATTAATTTCTTTTGAAGTGAAGTCTGCTGTGTGCCAG 2760
D 2701 TGTCTTATTAATTTTCTTAATTAATTTCTTTTGAAGTGAAGTCTGCTGTGTGCCAG 2760
Q 2761 GCTGAGTACAGTGTGTATATCCAGCTCACTGCAACCTCCGCTCTGGGTTCAATGA 2820
D 2761 GCTGAGTACAGTGTGTATATCCAGCTCACTGCAACCTCCGCTCTGGGTTCAATGA 2820
Q 2821 TTCTCTGCTCAGCTTCTTAAGTACTGAGGTTTACAGGACCCACACATGCCCCAGCT 2880
D 2821 TTCTCTGCTCAGCTTCTTAAGTACTGAGGTTTACAGGACCCACACATGCCCCAGCT 2880
Q 2881 AATTTTGTATTTTAAATAGACAGGGGTTTGGCCATTTGGCCAGGCTGTGTAATC 2940
D 2881 AATTTTGTATTTTAAATAGACAGGGGTTTGGCCATTTGGCCAGGCTGTGTAATC 2940
Q 2941 CTGACGTCAAGTATCTGCTGCTGCTGCTCCCAATACAGGATGAACCATGCAACCA 3000
D 2941 CTGACGTCAAGTATCTGCTGCTGCTGCTCCCAATACAGGATGAACCATGCAACCA 3000
Q 3001 CCTACTTAATTTTCAATGCTATGATGATTAAGAGATTTTTCATTTTCCATGACAT 3060
D 3001 CCTACTTAATTTTCAATGCTATGATGATTAAGAGATTTTTCATTTTCCATGACAT 3060
Q 3061 TTTTCTCTCTGCAATGCTTAAAGTACTGTGTGTTTCCCTTTTGGGCAAGACAGAT 3120
D 3061 TTTTCTCTCTGCAATGCTTAAAGTACTGTGTGTTTCCCTTTTGGGCAAGACAGAT 3120
Q 3121 CATTAATATTTCTGATCAATTTTCTTATCAAGAGATATATCAAGTGTGTCTCATAGA 3180
D 3121 CATTAATATTTCTGATCAATTTTCTTATCAAGAGATATATCAAGTGTGTCTCATAGA 3180
Q 3181 ACTGCTGATTCATTTATGTTTCTGATTCATCTGTGTCCCTTCATCTTGAC 3240
D 3181 ACTGCTGATTCATTTATGTTTCTGATTCATCTGTGTCCCTTCATCTTGAC 3240
Q 3241 TCTTTGTATTTCACTGAATTTCAAACTTTGTCAAGAGAAAATAATGAGAGTCTGAG 3300
D 3241 TCTTTGTATTTCACTGAATTTCAAACTTTGTCAAGAGAAAATAATGAGAGTCTGAG 3300
Q 3301 GAAAAATTAATTAATTAAGAACAGCTTTTGGGCGCGGAATTC 3345
D 3301 GAAAAATTAATTAATTAAGAACAGCTTTTGGGCGCGGAATTC 3345

RESULT 5
ABL64424
ID ABL64424 standard; DNA; 3345 BP.
XX
AC ABL64424;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2761.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209472P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234099P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234503P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235083P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237608P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Sopet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.

XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX
PS Claim 1; SEQ ID 2761; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL64
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Query Match 100.0%; Score 3345; DB 24; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGTCTGAGCAGCATGATGAGAAAGACCTTTAAACCATTTGTGACTTA 60
DB 1 GAATTCGGTCTGAGCAGCATGATGAGAAAGACCTTTAAACCATTTGTGACTTA 60

QY 61 CAGAAAGAAATTTGAATTAAGAAAGAACTATGATCTTCAGGCCATCTTCCTGCTGT 120
DB 61 CAGAAAGAAATTTGAATTAAGAAAGAACTATGATCTTCAGGCCATCTTCCTGCTGT 120

QY 121 CTCTTATGCTTTATTTGGCAATGATGATGCGCAAGAGGGAGAAAGTTAGTGACCCCTG 180
DB 121 CTCTTATGCTTTATTTGGCAATGATGATGCGCAAGAGGGAGAAAGTTAGTGACCCCTG 180

QY 181 AAACCATGACATTTCTATTATTTATGAGGCAAGAACCGAGTCAATATATTCAGATT 240
DB 181 AAACCATGACATTTCTATTATTTATGAGGCAAGAACCGAGTCAATATATTCAGATT 240

QY 241 AAGGCCAATCTCTCTGCTGACTTTGAACTAATCTGGGAGACAGCAACATATTTTGTG 300
DB 241 AAGGCCAATCTCTCTGCTGACTTTGAACTAATCTGGGAGACAGCAACATATTTTGTG 300

QY 301 ATGAAAGGAGGAGGACTTCTGTATTTACACAGAGCCTTGGACAGGAAACAAATCTACT 360
DB 301 ATGAAAGGAGGAGGACTTCTGTATTTACACAGAGCCTTGGACAGGAAACAAATCTACT 360

QY 361 CACAATTCAGAGTTGAGGAGCCCTGAGCGGTAATGAAATTAATGAGAGGCTCCAGTCCCT 420
DB 361 CACAATTCAGAGTTGAGGAGCCCTGAGCGGTAATGAAATTAATGAGAGGCTCCAGTCCCT 420

QY 421 ATCACCATAGAGTGAAGACATCAACGACATCGAACCCAGTTTCTCCAGTCAAGATAC 480
DB 421 ATCACCATAGAGTGAAGACATCAACGACATCGAACCCAGTTTCTCCAGTCAAGATAC 480

QY 481 GAAAGGCTAGTAAAGCAAGAACTCTCGCCGCAAGAAAGCCCTTTCTGTATGTCATGCCACA 540
DB 481 GAAAGGCTAGTAAAGCAAGAACTCTCGCCGCAAGAAAGCCCTTTCTGTATGTCATGCCACA 540

QY 541 GACCTGATATATCCGCACTCCCAATGAGGAGGAGTTTATTAACGATTTGATTCACACTT 600
DB 541 GACCTGATATATCCGCACTCCCAATGAGGAGGAGTTTATTAACGATTTGATTCACACTT 600

QY 601 CCCATGATCAACAATGATGATCTTTCAGATCAACAACAAAGGAGGAGGATCTCTCTT 660
DB 601 CCCATGATCAACAATGATGATCTTTCAGATCAACAACAAAGGAGGAGGATCTCTCTT 660

Db 601 CCGATGATCAACAAATGTCATCTTACATTCATCAACAAACGAGGAGCCATCTCTT 660
Qy 661 ACCGAGAGGGATCTCAGAAATGAATCCTCTAAGAATCCTCTAATCTGATC 720
Db 661 ACCGAGAGGGATCTCAGAAATGAATCCTCTAAGAATCCTCTAATCTGATC 720
Qy 721 TCATGGAAGACATGAGGAGCCAGAGTAGAATTCCTTCACTGATACCACTCTGAT 780
Db 721 TCATGGAAGACATGAGGAGCCAGAGTAGAATTCCTTCACTGATACCACTCTGAT 780
Qy 781 ATCATAGTGAACAGAAATATTGGAAGCCCAAAACCTGAGAGATGAGTGAACCTCA 840
Db 781 ATCATAGTGAACAGAAATATTGGAAGCCCAAAACCTGAGAGATGAGTGAACCTCA 840
Qy 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAATAT 900
Db 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCCGGTGCAATAT 900
Qy 901 TCCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTTGACAGAGAGAT 960
Db 901 TCCTTAGTTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTTGACAGAGAGAT 960
Qy 961 ATTACGTGACAGCCCTTGAGACCGAGAGAAAGAGATGATGTTTTTATGACGTT 1020
Db 961 ATTACGTGACAGCCCTTGAGACCGAGAGAAAGAGATGATGTTTTTATGACGTT 1020
Qy 1021 GCAAAAGATAGTACGAGAAACCACTTTCATATCCGCTGAAATTCATGTAAGTAA 1080
Db 1021 GCAAAAGATAGTACGAGAAACCACTTTCATATCCGCTGAAATTCATGTAAGTAA 1080
Qy 1081 GATATTAATGAATATCCACCTACATGTCCTGACACAGTAAACGTAATTTGAGGTGAG 1140
Db 1081 GATATTAATGAATATCCACCTACATGTCCTGACACAGTAAACGTAATTTGAGGTGAG 1140
Qy 1141 AATGAAGAGTGGGTACAGATTCGAGACCCCTTACTGACATGACAGAGATGAAGAAAT 1200
Db 1141 AATGAAGAGTGGGTACAGATTCGAGACCCCTTACTGACATGACAGAGATGAAGAAAT 1200
Qy 1201 ACTGCCAAGTCTTCTAATCTAAGATTTGAGAGAACTCCCAAACTTCCATGAT 1260
Db 1201 ACTGCCAAGTCTTCTAATCTAAGATTTGAGAGAACTCCCAAACTTCCATGAT 1260
Qy 1261 GGAATCTTCTAATCTAAGATTTGAGAGAACTCCCAAACTTCCATGAT 1320
Db 1261 GGAATCTTCTAATCTAAGATTTGAGAGAACTCCCAAACTTCCATGAT 1320
Qy 1321 AAGCAAGATCTCTCTGATCAACTTAAAGATGAGTGTCTGACAAAGATTTCAAGACC 1380
Db 1321 AAGCAAGATCTCTCTGATCAACTTAAAGATGAGTGTCTGACAAAGATTTCAAGACC 1380
Qy 1381 CTTGTTTGTGCAAAATCAAGTATTAATGATATCAATGATCCCATCTTTGAAAA 1440
Db 1381 CTTGTTTGTGCAAAATCAAGTATTAATGATATCAATGATCCCATCTTTGAAAA 1440
Qy 1441 TCAGATTTATGAAACCTGACTCTTGTGAGAACAAACATTTGGTCTCAACATCTTA 1500
Db 1441 TCAGATTTATGAAACCTGACTCTTGTGAGAACAAACATTTGGTCTCAACATCTTA 1500
Qy 1501 ATCCAGGACCTGATGCTGATGAGCCATTTACTGGAGATTCTAAATCTGTATCATATC 1560
Db 1501 ATCCAGGACCTGATGCTGATGAGCCATTTACTGGAGATTCTAAATCTGTATCATATC 1560
Qy 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATCAACCGGA 1620
Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATCAACCGGA 1620
Qy 1621 TATGATTAATTAAGGCTCTGATTTTGAACAGACGCTGTTTCAACATTTGTTTC 1680
Db 1621 TATGATTAATTAAGGCTCTGATTTTGAACAGACGCTGTTTCAACATTTGTTTC 1680
Qy 1681 AAGGAGAAATCTGAGCTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1740
Db 1681 AAGGAGAAATCTGAGCTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1740

Qy 1741 AAGTCAAGCTTATTTGACAGATGATGAATGAAGACCTCAATTTTCCCAACGATTC 1800
Db 1741 AAGTCAAGCTTATTTGACAGATGATGAATGAAGACCTCAATTTTCCCAACGATTC 1800
Qy 1801 CAAGCGAAGTCAAGTGAAGATGATGATGAAGCACTAAATGAGGCAATGATGATGAT 1860
Db 1801 CAAGCGAAGTCAAGTGAAGATGATGATGAAGCACTAAATGAGGCAATGATGATGAT 1860
Qy 1861 GATCCGAAGTCTGACATTAAGCTATTCCTGAGGAGAGACCAAGAGTTGGCTTAA 1920
Db 1861 GATCCGAAGTCTGACATTAAGCTATTCCTGAGGAGAGACCAAGAGTTGGCTTAA 1920
Qy 1921 ATTGACCAAGTCTGAGATCTTTAGTGGTCTCAATTTGACAGAGAGCGAGAT 1980
Db 1921 ATTGACCAAGTCTGAGATCTTTAGTGGTCTCAATTTGACAGAGAGCGAGAT 1980
Qy 1981 CCAATGCGGTACAGTGGTGGCCACAGAGTAGGGGGGTCTTCTTAAGCTCTGTGTC 2040
Db 1981 CCAATGCGGTACAGTGGTGGCCACAGAGTAGGGGGGTCTTCTTAAGCTCTGTGTC 2040
Qy 2041 GAGTTCACCTGATTCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 GAGTTCACCTGATTCCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Qy 2101 ACGGCTGTTCTTCTGACATCCCTCAGTGCACCTGGAAGTCTCAATTTTGAAGTACT 2160
Db 2101 ACGGCTGTTCTTCTGACATCCCTCAGTGCACCTGGAAGTCTCAATTTTGAAGTACT 2160
Qy 2161 GATGATGATCAGACCTAATTTGAGGAGTCCCATTTTACATTTTCCCTGCGCAGTGAAG 2220
Db 2161 GATGATGATCAGACCTAATTTGAGGAGTCCCATTTTACATTTTCCCTGCGCAGTGAAG 2220
Qy 2221 TTAACAAAGACCTGGAAGTTTCCAAATCAATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTAACAAAGACCTGGAAGTTTCCAAATCAATGATGATGATGATGATGATGATGAT 2280
Qy 2281 CACACAGACTTTGAGGAGAGGGGTATGCTGATGATGATGATGATGATGATGATGAT 2340
Db 2281 CACACAGACTTTGAGGAGAGGGGTATGCTGATGATGATGATGATGATGATGATGAT 2340
Qy 2341 CCAACCTTGAAGGACATTTCTTACAGTATTCCTGACAGTGTGATGATGATGATGAT 2400
Db 2341 CCAACCTTGAAGGACATTTCTTACAGTATTCCTGACAGTGTGATGATGATGATGAT 2400
Qy 2401 TGTTCGCGCAGACAGTCAACAGCTGGGATACCACTGTGGGCAATGGATGATGAT 2460
Db 2401 TGTTCGCGCAGACAGTCAACAGCTGGGATACCACTGTGGGCAATGGATGATGAT 2460
Qy 2461 CTGCTGACCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 2461 CTGCTGACCACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Qy 2521 AAGGATTAAGGCAAGATATGTTGAAGTGTCAAGCATCTGAAGTCAAACTCTGAG 2580
Db 2521 AAGGATTAAGGCAAGATATGTTGAAGTGTCAAGCATCTGAAGTCAAACTCTGAG 2580
Qy 2581 AGCTGAATTTGAAGGAATGTTGAATTTATTAAGCAAGTGTATTTGACAAACCA 2640
Db 2581 AGCTGAATTTGAAGGAATGTTGAATTTATTAAGCAAGTGTATTTGACAAACCA 2640
Qy 2641 TCTCATCTTATTAATTTTCAATCACTGATGATGATGATGATGATGATGATGATGAT 2700
Db 2641 TCTCATCTTATTAATTTTCAATCACTGATGATGATGATGATGATGATGATGATGAT 2700
Qy 2701 TGTCTTATTAATTTTGTCAATATTTTCTTTTGAAGTGAAGTGTGCTGTGCTGAG 2760
Db 2701 TGTCTTATTAATTTTGTCAATATTTTCTTTTGAAGTGAAGTGTGCTGTGCTGAG 2760
Qy 2761 GCTGAGATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820
Db 2761 GCTGAGATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820

QY	2821	TTCTCCGCGCCAGAGTTCCTAAGTAGAGTGGGTTTACAGAGCACCCACACATGCGCAGCT	2880
Db	2821	TTCTCTGCGCCACAGTCTCTAAGTACTGGGTTTACAGAGCACCCACACATGCGCAGCT	2880
QY	2881	AATTTTGTATTTTAAATAGAGAGCGGGTTTCGCCATTTTGGCCAGGCTGATCTTGAATC	2940
Db	2881	AATTTTGTATTTTAAATAGAGAGCGGGTTTCGCCATTTTGGCCAGGCTGATCTTGAATC	2940
QY	2941	CTGACGTCAGATGATCTGCGCTGCGCTTGGTCTTCCCAATACAGGCATGAAACCTGACCCA	3000
Db	2941	CTGACGTCAGATGATCTGCGCTGCGCTTGGTCTTCCCAATACAGGCATGAAACCTGACCCA	3000
QY	3001	CCTACTTAGATATTTTCATGCTGTATAGACATTTAGAGAAATTTTCATTTTCCATGACAT	3060
Db	3001	CCTACTTAGATATTTTCATGCTGTATAGACATTTAGAGAAATTTTCATTTTCCATGACAT	3060
QY	3061	TTTTCCTCTCTGCAAAATGGCTTAGCTTACTGTGTGTTTCCCTTTTGGGGCAAGACAGCT	3120
Db	3061	TTTTCCTCTCTGCAAAATGGCTTAGCTTACTGTGTGTTTCCCTTTTGGGGCAAGACAGCT	3120
QY	3121	CATTAAATATTTCTGTACATTTTCTTTATCAGAGAAATATATCAGTGTGTCTCATAGA	3180
Db	3121	CATTAAATATTTCTGTACATTTTCTTTATCAGAGAAATATATCAGTGTGTCTCATAGA	3180
QY	3181	ACTGCGCTGATTCATTTATGTTTTTTCGAAATCCATCCTGTGTCCTTCATCCCTTGAC	3240
Db	3181	ACTGCGCTGATTCATTTATGTTTTTTCGAAATCCATCCTGTGTCCTTCATCCCTTGAC	3240
QY	3241	TCCTTTGGTATTTTCACTGAATTTCAACAATTTTGTACAGAGAAAGAAAGTGAAGACTCAG	3300
Db	3241	TCCTTTGGTATTTTCACTGAATTTCAACAATTTTGTACAGAGAGAAAGAAAGTGAAGACTCAG	3300
QY	3301	GAAGAAATTAATTAATTAAGAAACAGCCTTTTGGCGCCGCGAATTC	3345
Db	3301	GAAGAAATTAATTAATTAAGAAACAGCCTTTTGGCGCCGCGAATTC	3345
RESULT 6			
ID	ABN97312	standard; DNA; 3654 BP.	
XX	AC	ABN97312;	
XX	DT	13-AUG-2002 (first entry)	
XX	DE	Gene #3810 used to diagnose liver cancer.	
XX	KM	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
XX	KM	metastatic liver tumour; cytostatic; expression profile; disease state;	
XX	OS	disease progression; drug toxicity; drug efficacy; drug metabolism.	
XX	OS	Homo sapiens.	
XX	PN	W0200229103-A2.	
XX	PD	11-APR-2002.	
XX	PF	02-OCT-2001; 2001WO-US30589.	
XX	PR	02-OCT-2000; 2000US-237054P.	
XX	PA	(GENE-) GENE LOGIC INC.	
XX	PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
XX	PI	WPI; 2002-426119/45.	
XX	PT	Diagnosing and detecting the progression of liver cancer,	
XX	PT	hepatocellular carcinoma or metastatic liver tumor in a patient,	
XX	PT	involves detecting the level of expression of two or more genes in a	
XX	PT	liver tissue sample -	
PS	Claim 1; SEQ ID NO 3810; 296bp; English.		

Query Match	Local Similarity	Score	DB	Length
Beet	Local Similarity	99.8%	Pred. No. 0	Matches 3519, Conservative 0, Mismatches 7, Indels 0, Gaps 0
Sequence	3654 BP, 1052 A, 804 C, 777 G, 1021 T, 0 other;			
8	GTCTCGACCACTGATGGAAGAAAAGACCTTTTAAACCAACATTTGTGACTTACAGAAAG	67		
13	GTCTCGACCACTGATGGAAGAAAAGACCTTTTAAACCAACATTTGTGACTTACAGAAAG	72		
68	GAATTGTAATAAGAAAAGCTATGATCTTCAGGCCCATCTTCACCTCGTGTCTTCTTA	127		
73	GAATTGTAATAAGAAAAGCTATGATCTTCAGGCCCATCTTCACCTCGTGTCTTCTTA	132		
128	TGCTTTATTTGGCAACGTGATATGAGCCAAAGAGGGAGTTTAAAGTGAACCCCTGAACCCA	187		
133	TGCTTTATTTGGCAACGTGATATGAGCCAAAGAGGGAGTTTAAAGTGAACCCCTGAACCCA	192		
188	TGACATTTTCTATTATTTATGAGAGCCCAAGAACCGAGTCAATATATATTCAGTTTAAAGCCA	247		
193	TGACATTTTCTATTATTTATGAGAGCCCAAGAACCGAGTCAATATATATTCAGTTTAAAGCCA	252		
248	ATCTCTCTGCTGTGACTTTTGAATCTAAGTGGGAGACAGACAAATATTTGTGATATGAAAC	307		
253	ATCTCTCTGCTGTGACTTTTGAATCTAAGTGGGAGACAGACAAATATTTGTGATATGAAAC	312		
308	GGGAGAGGACCTCTGATATTAACAAGAGGCTTGGACAGAGGAAAACAAGATCTACTCAACATC	367		
313	GGGAGAGGACCTCTGATATTAACAAGAGGCTTGGACAGAGGAAAACAAGATCTACTCAACATC	372		
358	TCCAGGTTGACAGCCCTGAGACGCTAATGGAATTAATAGTAGAGGGTCACTGCCCTATACCA	427		
373	TCCAGGTTGACAGCCCTGAGACGCTAATGGAATTAATAGTAGAGGGTCACTGCCCTATACCA	432		
428	TAGAAGTGAAGACATCAACGACATCGAACCCAGCTTTCTCAGTCAAAAGTACGAAGGCT	487		
433	TAGAAGTGAAGACATCAACGACATCGAACCCAGCTTTCTCAGTCAAAAGTACGAAGGCT	492		
488	CAGTAAAGGACAACTCTGCGCCAGAGAAAGCCCTTTTGATATGTCATATGACACAGACCTGG	547		
493	CAGTAAAGGACAACTCTGCGCCAGAGAAAGCCCTTTTGATATGTCATATGACACAGACCTGG	552		
548	ATGATCTCGGACCACTCCCAATGGCCAGCTTTATTAACAAGTGTCAATCCAGCTTCCCATGA	607		
553	ATGATCTCGGACCACTCCCAATGGCCAGCTTTATTAACAAGTGTCAATCCAGCTTCCCATGA	612		
608	TCAACAATGTCAATGACTTTCAGATCAACAACAAACGGAGCCATCTCTTTAACCAGAG	667		
613	TCAACAATGTCAATGACTTTCAGATCAACAACAAACGGAGCCATCTCTTTAACCAGAG	672		
668	AGGAAATTCAGAAATTTGAATCTCTGCTAAGAAATCTTCTCTATATATCTGGAGATCTCAGTA	727		
673	AGGAAATTCAGAAATTTGAATCTCTGCTAAGAAATCTTCTCTATATATCTGGAGATCTCAGTA	732		
728	AGGAATGGGAGGCAAGTAGGAATTCCTTCAAGTATCCACATCTGTGATATCATAG	787		

Db 733 AGGACATGGAGGCCAGAGTGAATTCCTTCACTGATACCAATCTGTGATATATCAAG 792
Qy 788 TGAAGAGAAATATTTGGAAAGCACAAACCTGTGAGATGTGGAAAACTCAACTGATC 847
Db 793 TGACAGAGAAATATTTGGAAAGCACAAACCTGTGAGATGTGGAAAACTCAACTGATC 852
Qy 848 CTGACCCCATCAAAATCACTCAGGTGGGTGGAAATGATCCCGGTGCAAAATATTCCTTAC 907
Db 853 CTGACCCCATCAAAATCACTCAGGTGGGTGGAAATGATCCCGGTGCAAAATATTCCTTAC 912
Qy 908 TTGACAAAGAGAGTGGCCGAAGTCCCATTTTCAATGACAGAGAGAGATATTTACG 967
Db 913 TTGACAAAGAGAGTGGCCGAAGTCCCATTTTCAATGACAGAGAGAGATATTTACG 972
Qy 968 TGACTGAGCCCTTGGACCGAGAGAAAGATGCAATGTTTTTTATGACAGTTGCAAAGG 1027
Db 973 TGACTGAGCCCTTGGACCGAGAGAAAGATGCAATGTTTTTTATGACAGTTGCAAAGG 1032
Qy 1028 ATGAGTACGGAAAAACCTTTCATATCCGCTGGAAATTCATGTAAAAGTTAAAGATATTA 1087
Db 1033 ATGAGTACGGAAAAACCTTTCATATCCGCTGGAAATTCATGTAAAAGTTAAAGATATTA 1092
Qy 1088 ATGATATATCCACTACATGTCCTGTCACACAGTAAACCGTATTTGAGTCCAGAGATGAAAC 1147
Db 1093 ATGATATATCCACTACATGTCCTGTCACACAGTAAACCGTATTTGAGTCCAGAGATGAAAC 1152
Qy 1148 GACTGGGTAAACAGTATCGGAGACCCTTATCTGACATGACAGAGATGAAAGAAATATCTGCCA 1207
Db 1153 GACTGGGTAAACAGTATCGGAGACCCTTATCTGACATGACAGAGATGAAAGAAATATCTGCCA 1212
Qy 1208 ACAGTTTTTCTAAACTACAGAGATGTGGAGCAAACTCCCAAACTTCCATGATGATGACTCT 1267
Db 1213 ACAGTTTTTCTAAACTACAGAGATGTGGAGCAAACTCCCAAACTTCCATGATGATGACTCT 1272
Qy 1268 TCCATATCCAAACCTATGCTGGAATGTACAGTTAGTAAACAGTCTTGTAAAGAGACAG 1327
Db 1273 TCCATATCCAAACCTATGCTGGAATGTACAGTTAGTAAACAGTCTTGTAAAGAGACAG 1332
Qy 1328 ATACTCCTCAGTACAACTTAACAGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1387
Db 1333 ATACTCCTCAGTACAACTTAACAGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTT 1392
Qy 1388 TTGTGCAAAATCAACGTTATTTGATATCAATGATCAGATCCCATCTTTGAAAAATCAAGTT 1447
Db 1393 TTGTGCAAAATCAACGTTATTTGATATCAATGATCAGATCCCATCTTTGAAAAATCAAGTT 1452
Qy 1448 ATGGAACCTGATCTTTGCTGGAAGACAAACATTGGGTCCACATCTTTAACATCCATCCAG 1507
Db 1453 ATGGAACCTGATCTTTGCTGGAAGACAAACATTGGGTCCACATCTTTAACATCCATCCAG 1512
Qy 1508 CCACTGATGCTGATGAGCCATTTACTGGGAGTTCTAAAATTTGTAATTCATATCATTAAGG 1567
Db 1513 CCACTGATGCTGATGAGCCATTTACTGGGAGTTCTAAAATTTGTAATTCATATCATTAAGG 1572
Qy 1568 GAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATACCAACACCGATATGTCA 1627
Db 1573 GAGACAGTGAAGGAGCCCTGGGGGTTGACACAGATCCCATACCAACACCGATATGTCA 1632
Qy 1628 TAATTTAAAAAGCTCTTGAATTTGAAAACAGACGCTGTTTCCACATGTGTGTTCAAGCAG 1687
Db 1633 TAATTTAAAAAGCTCTTGAATTTGAAAACAGACGCTGTTTCCACATGTGTGTTCAAGCAG 1692
Qy 1688 AAAATCCTGAGCCTAGTGTGTTGTTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1747
Db 1693 AAAATCCTGAGCCTAGTGTGTTGTTGTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1752
Qy 1748 CGCTTATTTGACAGATGTGAATGAAGACCTCAATTTTCCCAACAGTATTTCAAGGCA 1807
Db 1753 CGCTTATTTGACAGATGTGAATGAAGACCTCAATTTTCCCAACAGTATTTCAAGGCA 1812
Qy 1808 AAGTCACTGAGAGTATGATCTATATGAGCACTTAAGTGGGCAATGTGACTGCCAAGATCCAG 1867
Db 1813 AAGTCACTGAGAGTATGATCTATATGAGCACTTAAGTGGGCAATGTGACTGCCAAGATCCAG 1872

Qy 1868 AAGTCTGACATAAAGCTATATCTAGAGGGGAGACACAAGAGGTTGGCTTAAATTTGACC 1927
Db 1873 AAGTCTGACATAAAGCTATATCTAGAGGGGAGACACAAGAGGTTGGCTTAAATTTGACC 1932
Qy 1928 ACGTGACTGGTGAATCTTTAGTGTGCTCCATTTGACAGAGAGACCGAAGTCCATATTC 1987
Db 1933 ACGTGACTGGTGAATCTTTAGTGTGCTCCATTTGACAGAGAGACCGAAGTCCATATTC 1992
Qy 1988 GGGTACAGTGTGGGACAGAGATGAGGGGGTCTTCTTAAGTCTGTGTCAAGATTC 2047
Db 1993 GGGTACAGTGTGGGACAGAGATGAGGGGGTCTTCTTGAAGTCTGTGTCAAGATTC 2052
Qy 2048 ACCGTATCCCTTATGATGTGAATGAACAACCTCCAGAGCTAGGCAAGGACCTACAGGGGCT 2107
Db 2053 ACCGTATCCCTTATGATGTGAATGAACAACCTCCAGAGCTAGGCAAGGACCTACAGGGGCT 2112
Qy 2108 TGTCTTGTGCCATCCCTCAGTGACCTGGAAGTCTCATTTTGAAGGCTACTGATATG 2167
Db 2113 TGTCTTGTGCCATCCCTCAGTGACCTGGAAGTCTCATTTTGAAGGCTACTGATATG 2172
Qy 2168 ATCAGACTTATTTGGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGTTTCAAA 2227
Db 2173 ATCAGACTTATTTGGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGTTTCAAA 2232
Qy 2228 ACGACTGGGAAGTTTCCAAATCAATGTACTATGCCCCGACTGTCTACAGGACACAG 2287
Db 2233 ACGACTGGGAAGTTTCCAAATCAATGTACTATGCCCCGACTGTCTACAGGACACAG 2292
Qy 2288 ACTTTGAGAGAGAGGGGATATGTCCTTGTATCCGATCAATGATGAGGGGTCCGCCACCT 2347
Db 2293 ACTTTGAGAGAGAGGGGATATGTCCTTGTATCCGATCAATGATGAGGGGTCCGCCACCT 2352
Qy 2348 TGGAGGCAATGTTTCTTTTACAGTTACATTTCTGCAATGTGTGGAAGAGATGTTTCC 2407
Db 2353 TGGAGGCAATGTTTCTTTTACAGTTACATTTCTGCAATGTGTGGAAGAGATGTTTCC 2412
Qy 2408 GGCAGCAGGTACACAGACTGGGATACCACTGTGGGCAATGGCAGTGTATATGCTGTA 2467
Db 2413 GGCAGCAGGTACACAGACTGGGATACCACTGTGGGCAATGGCAGTGTATATGCTGTA 2472
Qy 2468 CCAACCTTCTGATGTGTATATTTTATAGCAAGTGTGTTTATCCGATTAAGAAAGATA 2527
Db 2473 CCAACCTTCTGATGTGTATATTTTATAGCAAGTGTGTTTATCCGATTAAGAAAGATA 2532
Qy 2528 AAGGCAAGATATGTTGAAAGTGTCAAGCATGTGAAGTCAAACTCTGAGAAAGCTGAA 2587
Db 2533 AAGGCAAGATATGTTGAAAGTGTCAAGCATGTGAAGTCAAACTCTGAGAAAGCTGAA 2592
Qy 2588 TTTGAAAAGGATGTTTGAATTTATATAGCAAGTGTGTTTATCCGATTAAGAAAGATA 2647
Db 2593 TTTGAAAAGGATGTTTGAATTTATATAGCAAGTGTGTTTATCCGATTAAGAAAGATA 2652
Qy 2648 CTATTAATTTTATCATCAAGTGCATTATATTTTAAACAGATATTCCTCTTGTGCTT 2707
Db 2653 CTATTAATTTTATCATCAAGTGCATTATATTTTAAACAGATATTCCTCTTGTGCTT 2712
Qy 2708 TAATATTTGCTAAATATTTTCTTTTGAAGTGAAGTCTTGTGCTGCGCCAGGCTGAG 2767
Db 2713 TAATATTTGCTAAATATTTTCTTTTGAAGTGAAGTCTTGTGCTGCGCCAGGCTGAG 2772
Qy 2768 TACAGTGTGTGATCCAGCTCAGTGAACCTCCGCTCTCTGGGTTCACATATCTCTCT 2827
Db 2773 TACAGTGTGTGATCCAGCTCAGTGAACCTCCGCTCTCTGGGTTCACATATCTCTCT 2832
Qy 2828 GCTCAGCTCTCTAAGTGTGAGTGTGTTTACAGGACCCACACACATGAGCCAGTAATTTT 2887
Db 2833 GCTCAGCTCTCTAAGTGTGAGTGTGTTTACAGGACCCACACACATGAGCCAGTAATTTT 2892
Qy 2888 GTATTTTAAATAGACAGGGGTTTGCATTTTGGCAGGCTGTGTTGAATCTCTGACGT 2947
Db 2893 GTATTTTAAATAGACAGGGGTTTGCATTTTGGCAGGCTGTGTTGAATCTCTGACGT 2952

Db 193 TGACATTTTCTATTATGAAAGGCCAAGAACCGAGTCAAAATTATATCCAGTTTAAAGGCCA 252
Qy 248 ATCTCTGCTGTGACTTTTGAACATACTGGGAGACAGACAACATATTTGTATAGAAC 307
Db 253 ATCTCTGCTGTGACTTTTGAACATACTGGGAGACAGACAACATATTTGTATAGAAC 312
Qy 308 GGGAGGACCTTCTGTATTATCAACAGAGCTTTGACAGAGAAAACAAGATCTACACATC 367
Db 313 GGGAGGACCTTCTGTATTATCAACAGAGCTTTGACAGAGAAAACAAGATCTACACATC 372
Qy 368 TCCAGGTTGACGCCCTGAGACGCTAATGGAATTATATGAGAGGCTCCAGTCCCTATACCA 427
Db 373 TCCAGGTTGACGCCCTGAGACGCTAATGGAATTATATGAGAGGCTCCAGTCCCTATACCA 432
Qy 428 TAGAAGTGAAGACATCAACGACATCGACCCAGTTCCTCAGTCAAAAGTACGAGGCT 487
Db 433 TAAAGTGAAGACATCAACGACATCGACCCAGTTCCTCAGTCAAAAGTACGAGGCT 492
Qy 488 CAGTAAAGGACAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCAATGCCACAGACCTGG 547
Db 493 CAGTAAAGGACAACTCTGCGCCAGAGAAAGCCCTTCTGTATGTCAATGCCACAGACCTGG 552
Qy 548 ATGATCCGGCCACTCCCAATGSCAGCTTTATACAGATTCAGATTCAGCTTCCCATGA 607
Db 553 ATGATCCGGCCACTCCCAATGSCAGCTTTATACAGATTCAGATTCAGCTTCCCATGA 612
Qy 608 TCAACAATGTCAATGACTTTTCAAGATCAACAAACAAAGGAGCCACTCTCTTACCCGAG 667
Db 613 TCAACAATGTCAATGACTTTTCAAGATCAACAAACAAAGGAGCCACTCTCTTACCCGAG 672
Qy 668 AGGAGTCTCAGAAATTGAATCTGTCTAAGAAATCTTCTCTATATCTGTGTATCTCAGTGA 727
Db 673 AGGAGTCTCAGAAATTGAATCTGTCTAAGAAATCTTCTCTATATCTGTGTATCTCAGTGA 732
Qy 728 AGGACATGGAGAGGCGACAGTGAAGAAATCTTCAAGTATACCAATCTGTGTATCTCAGTGA 787
Db 733 AGGACATGGAGAGGCGACAGTGAAGAAATCTTCAAGTATACCAATCTGTGTATCTCAGTGA 792
Qy 788 TGAACAGAAATATTTGAAAGACCAAAACCTGTGAGATGTGTGAAAGAACTCAACTGATC 847
Db 793 TGAACAGAAATATTTGAAAGACCAAAACCTGTGAGATGTGTGAAAGAACTCAACTGATC 852
Qy 848 CTGACCCCATCAAAATCACTCAGGTGCGGTGAATGAATCCCGGTGCAAAATATCTCTTAA 907
Db 853 CTGACCCCATCAAAATCACTCAGGTGCGGTGAATGAATCCCGGTGCAAAATATCTCTTAA 912
Qy 908 TTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTGACAGAGAAAGATATTTAGC 967
Db 913 TTGACAAAGAGAGCTGCCAAGATTCCTCATTTTCAATTGACAGAGAAAGATATTTAGC 972
Qy 968 TGACTCAGCCCTTGACCGAGAGAAAGATGATATGTTTTTATGAGTTGCAAGG 1027
Db 973 TGACTCAGCCCTTGACCGAGAGAAAGATGATATGTTTTTATGAGTTGCAAGG 1032
Qy 1028 ATGAGTACGAGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAAAGATATTA 1087
Db 1033 ATGAGTACGAGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAAAGATATTA 1092
Qy 1088 ATGATATTCACCTACATGTCCGTACACAGTAACCGTATTTAGGTCCAGAGATATAC 1147
Db 1093 ATGATATTCACCTACATGTCCGTACACAGTAACCGTATTTAGGTCCAGAGATATAC 1152
Qy 1148 GACTGGGTACAGATGCGGACCCCTTATGACATGACAGAGATGAAGAAATATCTGCCA 1207
Db 1153 GACTGGGTACAGATGCGGACCCCTTATGACATGACAGAGATGAAGAAATATCTGCCA 1212
Qy 1208 ACAGTTTTCTAAACCTACAGAGATGTGAGCAAACTCCCAACTTCCATGATGATCTCT 1267
Db 1213 ACAGTTTTCTAAACCTACAGAGATGTGAGCAAACTCCCAACTTCCATGATGATCTCT 1272
Qy 1268 TCTTAATCCAACTTATGCTGGAATGTATTAAGTTAAGCTTCTTGAAGAGCAAG 1327
Db 1273 TCTTAATCCAACTTATGCTGGAATGTATTAAGTTAAGCTTCTTGAAGAGCAAG 1332

Qy 1328 ATACTCTCAGTACAACTTAACATAGAGGTGTCTGACAAAGATTTCAAGACCCTTTGT 1387
Db 1333 ATACTCTCAGTACAACTTAACATAGAGGTGTCTGACAAAGATTTCAAGACCCTTTGT 1392
Qy 1388 TTGTGCAATCAACGTTATTTGATATCAATGATCAGATGCCCATCTTTGAAATAACAGATT 1447
Db 1393 TTGTGCAATCAACGTTATTTGATATCAATGATCAGATGCCCATCTTTGAAATAACAGATT 1452
Qy 1448 ATGAAACCTGACTCTTGTGGAAGACAAACATTTGGGTCCACATCTTTAAACATCCAGG 1507
Db 1453 ATGAAACCTGACTCTTGTGGAAGACAAACATTTGGGTCCACATCTTTAAACATCCAGG 1512
Qy 1508 CCACTGATGCTGATGAGCCATTACTGGAGGTTCTTAAATTTCTGTATCATATCAATAAGG 1567
Db 1513 CCACTGATGCTGATGAGCCATTACTGGAGGTTCTTAAATTTCTGTATCATATCAATAAGG 1572
Qy 1568 GAGACAGTGAAGGACGCTGGGGGTTGACACAGATCCCATACCAACCCGATATGTCA 1627
Db 1573 GAGACAGTGAAGGACGCTGGGGGTTGACACAGATCCCATACCAACCCGATATGTCA 1632
Qy 1628 TAAATTAAAGCCCTTGTGATTTTGAACACAGACCTGTTTCCACATGTGTGAAAGCAG 1687
Db 1633 TAAATTAAAGCCCTTGTGATTTTGAACACAGACCTGTTTCCACATGTGTGAAAGCAG 1692
Qy 1688 AAAATCCTGAGCCCTAGTGTGTGTGTAAGTACAAATGCAAGTCTTTTGCCAAAGTTCA 1747
Db 1693 AAAATCCTGAGCCCTAGTGTGTGTGTAAGTACAAATGCAAGTCTTTTGCCAAAGTTCA 1752
Qy 1748 CGCTTATTTGTGACAGATGTGAATGAAGACCTCAATTTTCCAAACGATATCCAAAGCA 1807
Db 1753 CGCTTATTTGTGACAGATGTGAATGAAGACCTCAATTTTCCAAACGATATCCAAAGCA 1812
Qy 1808 AAGTCAGTGAAGATGTATAGCTATATAGGCACTTAAATGGGCAATGTGACTGCCAAGATCCAG 1867
Db 1813 AAGTCAGTGAAGATGTATAGCTATATAGGCACTTAAATGGGCAATGTGACTGCCAAGATCCAG 1872
Qy 1868 AAGTCTGACATATAGCTATCTGAGGGGAGACCAAGAGGTTGGCTTAAATGAGC 1927
Db 1873 AAGTCTGACATATAGCTATCTGAGGGGAGACCAAGAGGTTGGCTTAAATGAGC 1932
Qy 1928 ACGTGACTGTGAGATCTTATAGTGTGCTCATTTGACAGAGAGCCGAAGTCCATATC 1987
Db 1933 ACGTGACTGTGAGATCTTATAGTGTGCTCATTTGACAGAGAGCCGAAGTCCATATC 1992
Qy 1988 GGGTACAGTGTGTGGCCACAGAAATGTAGGGGGGCTTCTTAACTCTGTCTCAGAGTTCC 2047
Db 1993 GGGTACAGTGTGTGGCCACAGAAATGTAGGGGGGCTTCTTAACTCTGTCTCAGAGTTCC 2052
Qy 2048 ACCGTATCCCTTATGAGATGTGAATGACAAACCTCCAGGCTAGGCAAGACTAACCGGGCT 2107
Db 2053 ACCGTATCCCTTATGAGATGTGAATGACAAACCTCCAGGCTAGGCAAGACTAACCGGGCT 2112
Qy 2108 TGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTCAATTTTCAAGGCTACGATGATG 2167
Db 2113 TGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTCAATTTTCAAGGCTACGATGATG 2172
Qy 2168 ATCAGCACTTATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAAA 2227
Db 2173 ATCAGCACTTATTTGGGGTCCCATTTTACATTTTCCCTGGCAGTGAAGCTTACAAA 2232
Qy 2228 ACGACTGGGAAGTTTCCAAATCAATGTATCAATCCCGACCTGCTTACACAGCACAAG 2287
Db 2233 ACGACTGGGAAGTTTCCAAATCAATGTATCAATCCCGACCTGCTTACACAGCACAAG 2292
Qy 2288 ACTTTGAGAGAGGGGCTATGTCTGTGATCCGATCAATGATGGGGGTGGCCACCT 2347
Db 2293 AGTTTGAAGAGAGGGGATGTGTCTGTATCCGATCAATATATGGGGGTGGCCACCT 2352
Qy 2348 TGAAGGCAATGTTTCTTTTACAGATTACATCTGCACTGTGTGGAAGAAAGTTGTTCC 2407
Db 2353 TGAAGGCAATGTTTCTTTTACAGATTACATCTGCACTGTGTGGAAGAAAGTTGTTCC 2412

QY 2408 GGGCAGCAGGTCAACAGCTGGGATACCACTGTGGGCAATGGAGTTGGTATCTGCTGA 2467
 |||||
 DB 2413 GGGCAGCAGGTCAACAGCTGGGATACCACTGTGGGCAATGGAGTTGGTATCTGCTGA 2472
 |||||
 QY 2468 CCACCCCTTGGGATGTGATTAATTTAGCAGTGTGTTATCCGCATTAAGAGATA 2527
 |||||
 DB 2473 CCACCCCTTGGGATGTGATTAATTTAGCAGTGTGTTATCCGCATTAAGAGATA 2532
 |||||
 QY 2528 AAGCAAAAGATTAATTTGAAAGTGTCAAGCATCTGAAGTCAAACTCTGAAGAGCTGA 2587
 |||||
 DB 2533 AAGCAAAAGATTAATTTGAAAGTGTCAAGCATCTGAAGTCAAACTCTGAAGAGCTGA 2592
 |||||
 QY 2588 TTGAAAAGGAATTTGAAATTTATATAGCAAGTCAATTTGCAAGCAACCATCTCATC 2647
 |||||
 DB 2593 TTGAAAAGGAATTTGAAATTTATATAGCAAGTCAATTTGCAAGCAACCATCTCATC 2652
 |||||
 QY 2648 CTATTACTTTTCATCTCAAGCTGCACTTAATTTTAAAGATATTCCTCTGTGCTT 2707
 |||||
 DB 2653 CTATTACTTTTCATCTCAAGCTGCACTTAATTTTAAAGATATTCCTCTGTGCTT 2712
 |||||
 QY 2708 TAAATTTTGTAAATATTTCTTTTGAAGTGAAGTCTGTGCTGTGCGCCAGGCTGAG 2767
 |||||
 DB 2713 TAAATTTTGTAAATATTTCTTTTGAAGTGAAGTCTGTGCTGTGCGCCAGGCTGAG 2772
 |||||
 QY 2768 TACAGTGTGTGATCCCACTCACTGCAACTCCGCTCCGCGGTTGCAATGATTCCT 2827
 |||||
 DB 2773 TACAGTGTGTGATCCCACTCACTGCAACTCCGCTCCGCGGTTGCAATGATTCCT 2832
 |||||
 QY 2828 GCGCTCAGCTTCCTTAAGTGTGGTTTAAAGGACCAACCAACATGCGCAAGCTAATTTT 2887
 |||||
 DB 2833 GCGCTCAGCTTCCTTAAGTGTGGTTTAAAGGACCAACCAACATGCGCAAGCTAATTTT 2892
 |||||
 QY 2888 GATTTTTTAAATAGAGACGGGGTTTCCCAATTTGGCCAGGCTGTCTTGAATCTCTGACGT 2947
 |||||
 DB 2893 GATTTTTTAAATAGAGACGGGGTTTCCCAATTTGGCCAGGCTGTCTTGAATCTCTGACGT 2952
 |||||
 QY 2948 CAAATGATCTGCGCTGCTTGTGTCTCCCAATTCAGGCAATGAAACAATGCCAACCCTACT 3007
 |||||
 DB 2953 CAAATGATCTGCGCTGCTTGTGTCTCCCAATTCAGGCAATGAAACAATGCCAACCCTACT 3012
 |||||
 QY 3008 AGATATTTCATGTGCTATAGACATTAAGAGATTTTTCATTTTCCATGACATTTTCT 3067
 |||||
 DB 3013 AGATATTTCATGTGCTATAGACATTAAGAGATTTTTCATTTTCCATGACATTTTCT 3072
 |||||
 QY 3068 CTCTGCAAAATGGCTTAACTATGTGTTTTCCCTTTTGGGCAAGACACTATTAA 3127
 |||||
 DB 3073 CTCTGCAAAATGGCTTAACTATGTGTTTTCCCTTTTGGGCAAGACACTATTAA 3132
 |||||
 QY 3128 TATCTGTACATTTTCTTTCTTATCAAGAGATATATCAAGTGTGCTCATGAACTGCT 3187
 |||||
 DB 3133 TATCTGTACATTTTCTTTCTTATCAAGAGATATATCAAGTGTGCTCATGAACTGCT 3192
 |||||
 QY 3188 GGATTTCAATTTATTTTCTTTCTGATTCATCTGTGCTCCCTTCACTCTTGG 3247
 |||||
 DB 3193 GGATTTCAATTTATTTTCTTTCTGATTCATCTGTGCTCCCTTCACTCTTGG 3252
 |||||
 QY 3248 GTATTTCACGTAATTTCAACAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3307
 |||||
 DB 3253 GTATTTCACGTAATTTCAACAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
 |||||
 QY 3308 AAAATTAATTAAG 3333
 |||||
 DB 3313 AAAATTAATTAAG 3338
 |||||

DE Oesophagus cancer related gene sequence SEQ ID NO:6004.
 XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
 XX gene; ds.
 OS Homo sapiens.
 XX
 EN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 XX
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 20-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237117P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237358P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.
 XX
 XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set
 XX
 XX Claim 1; SEQ ID 6004; 44p; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical

CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (1)
CC comprises a sequence (S) selected from 8447 sequences (given in AB161664
CC to AB170110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity, (2) has cytostatic
CC activity and can be used in gene therapy, M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent, M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;

Query Match 99.1%; Score 3314.8; DB 24; Length 3654;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 8 GTCTCGACCACTGAATGGAAGAAAGACCTTTTACCACTTTTGACTTACAGAAAG 67
DB 13 GTCTCGACCACTGAATGGAAGAAAGACCTTTTAAACCACTTTTGACTTACAGAAAG 72
QY 68 GAATTTGAATTAAGAAAGAACTATGATCTTACAGGCCCATCTTCCCTGCTGCTTTA 127
DB 73 GAATTTGAATTAAGAAAGAACTATGATCTTACAGGCCCATCTTCCCTGCTGCTTTA 132
QY 128 TGCCTTAATTTGGCACTGATATGCGCAAGAGGGGAAGTTTATGAGACCCCTGAAACCA 187
DB 133 TGCCTTAATTTGGCACTGATATGCGCAAGAGGGGAAGTTTATGAGACCCCTGAAACCA 192
QY 188 TGACATTTTCTATTTATGAAGGCCCAAGAACCGAGTCAATTTATTTCACTTTAAGGCCA 247
DB 193 TGACATTTTCTATTTATGAAGGCCCAAGAACCGAGTCAATTTATTTCACTTTAAGGCCA 252
QY 248 ATCTCTCTGCTGATCTTTGAAGTAACTAGTGGGAGACAGACATATTTGTATGATGAAC 307
DB 253 ATCTCTCTGCTGATCTTTGAAGTAACTAGTGGGAGACAGACATATTTGTATGATGAAC 312
QY 308 GGGAGGGAATCTGTATTTACACAGAGCCTTGGACAGGGAACAAAGATCTTACACATC 367
DB 313 GGGAGGGAATCTGTATTTACACAGAGCCTTGGACAGGGAACAAAGATCTTACACATC 372
QY 368 TCCAGGTTGAGCCCTGAGAGCTTATGAGATTTATGAGAGGCTCCAGTCCCTATCCCA 427
DB 373 TCCAGGTTGAGCCCTGAGAGCTTATGAGATTTATGAGAGGCTCCAGTCCCTATCCCA 432
QY 428 TAGAAGTGAAGGACATCAACGACATGACCCAGTTTCTCCAGTCAAGTAGAAGGCT 487
DB 433 TAGAAGTGAAGGACATCAACGACATGACCCAGTTTCTCCAGTCAAGTAGAAGGCT 492
QY 488 CAGTAAAGGCAAACTCTCGCCAGAAAGCCTTCTGTATGTCATATGACCAAGACCTGG 547
DB 493 CAGTAAAGGCAAACTCTCGCCAGAAAGCCTTCTGTATGTCATATGACCAAGACCTGG 552
QY 548 ATATATCCGCGCACTCCCAATGCGCAAGCTTTATTAACAGATTTGATCCAGCTTCCATGA 607
DB 553 ATATATCCGCGCACTCCCAATGCGCAAGCTTTATTAACAGATTTGATCCAGCTTCCATGA 612
QY 608 TCAACAGATGTCATGATCTTCAATCAACAAAGAGGAGCATCTCTTACCCGAG 667
DB 613 TCAACAGATGTCATGATCTTCAATCAACAAAGAGGAGCATCTCTTACCCGAG 672
QY 668 AGGAGATCTCAGGAATGAATCTCTGTAAGATCTTCTATTAATCTGTATCTCAGTGA 727
DB 673 AGGAGATCTCAGGAATGAATCTCTGTAAGATCTTCTATTAATCTGTATCTCAGTGA 732
QY 728 AGGACATGGAGAGCCAGAGTGAATTTCTTCAAGTATACCAATCTGTGATATCATAG 787
DB 733 AGGACATGGAGAGCCAGAGTGAATTTCTTCAAGTATACCAATCTGTGATATCATAG 792

QY 788 TGACAGAGATATTTGGAAAAAGCACAAAACCTGTGAGATGTGAAAACTCACTGATC 847
DB 793 TGACAGAGATATTTGGAAAAAGCACAAAACCTGTGAGATGTGAAAACTCACTGATC 852
QY 848 CTACACCCCATCAAAATCTACTCAGTGGCGGTGAATATCCCGGTGCACAATATCTTTAG 907
DB 853 CTACACCCCATCAAAATCTACTCAGTGGCGGTGAATATCCCGGTGCACAATATCTTTAG 912
QY 908 TTGACAAAGAGAGCTGCCAAGATTCCTTCAATTTTGAACCGAAGAGATATTTAGC 967
DB 913 TTGACAAAGAGAGCTGCCAAGATTCCTTCAATTTTGAACCGAAGAGATATTTAGC 972
QY 968 TGACTAGCCCTTGGACCCGAGAAAGAAAGATCATATGTTTTTATGAGTGGCAAGG 1027
DB 973 TGACTAGCCCTTGGACCCGAGAAAGAAAGATCATATGTTTTTATGAGTGGCAAGG 1032
QY 1028 ATGAGTACGGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAAAGATATTA 1087
DB 1033 ATGAGTACGGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAAAGATATTA 1092
QY 1088 ATGATTAATCCACTATCATATGTCGTACCAAGTAAACCGTATTTGAGTCCAGAGAAATGAC 1147
DB 1093 ATGATTAATCCACTATCATATGTCGTACCAAGTAAACCGTATTTGAGTCCAGAGAAATGAC 1152
QY 1148 GACTGGGTAAACAGTATCGGGGACCTTACTGACATGACAGAGGATGAAGAAAAATACGCCA 1207
DB 1153 GACTGGGTAAACAGTATCGGGGACCTTACTGACATGACAGAGGATGAAGAAAAATACGCCA 1212
QY 1208 ACAGTTTTCTTAACTACAGAGATTTGTGAGCAAACTCCCAACTTCCATGATGATCTCT 1267
DB 1213 ACAGTTTTCTTAACTACAGAGATTTGTGAGCAAACTCCCAACTTCCATGATGATCTCT 1272
QY 1268 TCTTAATCCAAACCTATGCTGGAATGTTCAGTTAGCTTAAACGCTCTTGAAGAGCAAG 1327
DB 1273 TCTTAATCCAAACCTATGCTGGAATGTTCAGTTAGCTTAAACGCTCTTGAAGAGCAAG 1332
QY 1328 ATACTCCTCAGTCAACTTAAGAGATGAGGTGCTGACAAAGTTTCAAGACCTTTGTT 1387
DB 1333 ATACTCCTCAGTCAACTTAAGAGATGAGGTGCTGACAAAGTTTCAAGACCTTTGTT 1392
QY 1388 TTGTGCAATCAACGTTATTTGATATCAATGATATGCCCATCTTTGAAAAATACGATT 1447
DB 1393 TTGTGCAATCAACGTTATTTGATATCAATGATATGCCCATCTTTGAAAAATACGATT 1452
QY 1448 ATGGAACCTGACTCTTGTCTGAAGACAAACATTTGGTTCACATCTTAACTCCAGG 1507
DB 1453 ATGGAACCTGACTCTTGTCTGAAGACAAACATTTGGTTCACATCTTAACTCCAGG 1512
QY 1508 CCACTGATGCTGATGAGCCATTTACTGGAGTTCTTAAATTTGATATCATATCAATAAG 1567
DB 1513 CCACTGATGCTGATGAGCCATTTACTGGAGTTCTTAAATTTGATATCATATCAATAAG 1572
QY 1568 GAGACAGTGAAGGACGCTGTGGGGTTGACACAGATCCCATACCAACCCGATATATGTA 1627
DB 1573 GAGACAGTGAAGGACGCTGTGGGGTTGACACAGATCCCATACCAACCCGATATATGTA 1632
QY 1628 TAATTAAGAGCTCTTGAATTTTGAACAGACGCTGTTTCCAAATTTGTTTCAAGAG 1687
DB 1633 TAATTAAGAGCTCTTGAATTTTGAACAGACGCTGTTTCCAAATTTGTTTCAAGAG 1692
QY 1688 AAAATCCTGAGCCTCAGAGTTTGTGTGTAATGACAAATGCAAGTTCTTTGCAAGTTCA 1747
DB 1693 AAAATCCTGAGCCTCAGAGTTTGTGTGTAATGACAAATGCAAGTTCTTTGCAAGTTCA 1752
QY 1748 CGCTTAATGTGACAGATGTAATGAAGACCTCAATTTTCCCAACAGTATTCAGAGCA 1807
DB 1753 CGCTTAATGTGACAGATGTAATGAAGACCTCAATTTTCCCAACAGTATTCAGAGCA 1812
QY 1808 AAGTCAGTGAAGATGATGATTAAGGACCTTAAAGTGGCAATGTGATGCTGCCAAGATCCAG 1867
DB 1813 AAGTCAGTGAAGATGATGATTAAGGACCTTAAAGTGGCAATGTGATGCTGCCAAGATCCAG 1872

QY 1868 AAGCTGTGACATPAGCTAATCTACTGAGGGGAGACAGAGAGTGGCTTAAATTAAC 1927
 DB 1873 AAGGTCTGAGACATPAGCTAATCTACTGAGGGGAGACAGAGAGTGGCTTAAATTAAC 1932
 QY 1928 ACGTGAATGTGAGATCTTTAGTGTGCTCCATTGGAAGAAGAGCCGGAATGCATATC 1987
 DB 1933 ACGTGAATGTGAGATCTTTAGTGTGCTCCATTGGAAGAAGAGCCGGAATGCATATC 1992
 QY 1988 GGGTACAGATGTGTGAGGAGAGAGATAGGGGGGTCTTCTTAAGCTCTGTGTCAAGTTCC 2047
 DB 1993 GGGTACAGATGTGTGAGGAGAGAGATAGGGGGGTCTTCTTAAGCTCTGTGTCAAGTTCC 2052
 QY 2048 ACCGATCTTATGATGTGATGATGACAACTCCAGAGCTPAGCCAGAGATACAGAGGCT 2107
 DB 2053 ACCGATCTTATGATGTGATGATGACAACTCCAGAGCTPAGCCAGAGATACAGAGGCT 2112
 QY 2108 TGTTCCTGTGCAATCCCTCAGTGCACCTGAGAGTCAATTTTGAGGCTA CTGATGATG 2167
 DB 2113 TGTTCCTGTGCAATCCCTCAGTGCACCTGAGAGTCAATTTTGAGGCTA CTGATGATG 2172
 QY 2168 ATCAGCACTTATTTGGGGGTCCCATTTTACATTTTCCCTGCGAGAGTGAAGCTTACAA 2227
 DB 2173 ATCAGCACTTATTTGGGGGTCCCATTTTACATTTTCCCTGCGAGAGTGAAGCTTACAA 2232
 QY 2228 ACGACTGGAGAGTTTCCAAAATCAATGATGATGATGATGATGATGATGATGATGATG 2287
 DB 2233 ACGACTGGAGAGTTTCCAAAATCAATGATGATGATGATGATGATGATGATGATGATG 2292
 QY 2288 ACTTGAAGAGAGGAGGCTATGCTGCTTGAATGCGCATCAATGATGATGATGATGATG 2347
 DB 2293 AGTTGAAGAGAGGAGGATGCTGCTTGAATGCGCATCAATGATGATGATGATGATGATG 2352
 QY 2348 TGAAGAGCATTTTCTTTTACAGTTTACATCTGCAAGTGTGAGAGAGAGAGTTTCC 2407
 DB 2353 TGAAGAGCATTTTCTTTTACAGTTTACATCTGCAAGTGTGAGAGAGAGAGTTTCC 2412
 QY 2408 GGCCAGAGAGTCAACAGACTGGAGATACCACTGAGGAGATGAGAGTGTATCTGCTGA 2467
 DB 2413 GGCCAGAGAGTCAACAGACTGGAGATACCACTGAGGAGATGAGAGTGTATCTGCTGA 2472
 QY 2468 CCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2527
 DB 2473 CCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2532
 QY 2528 AAGCAAAGATTAATGTTGAAGTGTCAAGATCTGAAGTAACTCTGAGAGCTGAA 2587
 DB 2533 AAGCAAAGATTAATGTTGAAGTGTCAAGATCTGAAGTAACTCTGAGAGCTGAA 2592
 QY 2588 TTGAAAAGAGATGTTGAATTTATATAGCAAGTGTATTTGAGCAACATCTCATC 2647
 DB 2593 TTGAAAAGAGATGTTGAATTTATATAGCAAGTGTATTTGAGCAACATCTCATC 2652
 QY 2648 CTATTACTTTTCACTAAGCTGATTAATTTTAAACAGATATTTCCCTTGTCTT 2707
 DB 2653 CTATTACTTTTCACTAAGCTGATTAATTTTAAACAGATATTTCCCTTGTCTT 2712
 QY 2708 TAAATTTTGTAAATATTTCTTTTGAAGTGTGATCTGTGTGCGCCAGAGCTGAG 2767
 DB 2713 TAAATTTTGTAAATATTTCTTTTGAAGTGTGATCTGTGTGCGCCAGAGCTGAG 2772
 QY 2768 TACAGTGTGTGATCCAGCTCACTGCAACTCCGCTCTGAGGTTCAATGATTTCTCT 2827
 DB 2773 TACAGTGTGTGATCCAGCTCACTGCAACTCCGCTCTGAGGTTCAATGATTTCTCT 2832
 QY 2828 GCTCAGCTTCTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2887
 DB 2833 GCTCAGCTTCTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2892
 QY 2888 GTATTTTAAATAGAGAGGAGTGTGCTGATTTGAGGAGGCTGTGAACTCTGAGCT 2947
 DB 2893 GTATTTTAAATAGAGAGGAGTGTGCTGATTTGAGGAGGCTGTGAACTCTGAGCT 2952
 QY 2948 CAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3007

DB 2953 CAATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3012
 QY 3008 AGATATTTTCAATGTGCTATATGACATTTAGAGATTTTTCATTTTCCATATATTTTCC 3067
 DB 3013 AGATATTTTCAATGTGCTATATGACATTTAGAGATTTTTCATTTTCCATATATTTTCC 3072
 QY 3068 CTCTGCAAAATGGGTAGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3127
 DB 3073 CTCTGCAAAATGGGTAGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3132
 QY 3128 TATTTGTACATTTTCTTTTATCAAGAGATATATCAATGTGTGTGTGTGTGTGTGTGT 3187
 DB 3133 TATTTGTACATTTTCTTTTATCAAGAGATATATCAATGTGTGTGTGTGTGTGTGTGT 3192
 QY 3188 GATTCATTTATGATTTTCTTTTATCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3247
 DB 3193 GATTCATTTATGATTTTCTTTTATCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3252
 QY 3248 GATTTTCACTGATTTTCAAAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3307
 DB 3253 GATTTTCACTGATTTTCAAAATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
 QY 3308 AATTAATTAAG 3333
 DB 3313 AATTAATTAAG 3338
 RESULT 9
 ABL37430
 ID ABL37430 standard; cDNA; 523 BP.
 XX
 AC ABL37430;
 DT 08-APR-2002 (first entry)
 XX
 DE Human colon tumour antigen polynucleotide SEQ ID NO:1019.
 XX
 KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 KW colon tumour metastatic antigen; diagnosis; gene; ss.
 OS Homo sapiens.
 XX
 PN NC0200196388-A2.
 XX
 PD 20-DEC-2001.
 XX
 PF 08-JUN-2001; 2001MO-US18557.
 XX
 PR 09-JUN-2000; 2000US-210899P.
 XX
 PR 20-FEB-2001; 2001US-270216P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Jiang Y, Harlocker SL, Secrist H;
 XX
 DR WPI; 2002-114514/15.
 XX
 PT Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX
 PS Claim 1; SEQ ID 1019; 105bp; English.
 XX
 CC ABL36412 to ABL38645 represent human colon tumour cDNA clones (1)
 CC which were isolated from human colon tumour and colon metastatic tumour
 CC cDNA libraries. (1) have cytoskeletal activity and can be used in vaccine
 CC production. (1) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (1) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (1)
 CC can be used in the diagnosis of a colon tumour.

XX Sequence 523 BP; 133 A; 107 C; 119 G; 162 T; 2 other;
SQ

Query Match 15.4%; Score 514.2; DB 24; Length 523;
Best Local Similarity 99.4%; Pred. No. 4.5e-127;
Matches 516; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2257 ACTCATGCCGACTGTGTACACGACACAGACTTTGAGGAGGCGGTATGTCGCTTG 2316
DB 1 ACTCATGCCGACTGTGTACACGACACAGACTTTGAGGAGGCGGTATGTCGCTTG 60
QY 2317 ATCCGATCATATGAGGGGTCGCCACCTTGGAGGCAATGTTTCTTTACCAAGTTACA 2376
DB 61 ATCCGATCATATGAGGGGTCGCCACCTTGGAGGCAATGTTTCTTTACCAAGTTACA 120
QY 2377 TTCTGAGTGTGTGGAAGAGAGTGTTCGCGACAGAGTACACGACTGGGATPCCC 2436
DB 121 TTCTGAGTGTGTGGAAGAGAGTGTTCGCGACAGAGTACACGACTGGGATPCCC 180
QY 2437 ACTGTGGGCAATGGCAGTTGGTATATCTGTACACACCTTCTGGTATTTGATATTTTA 2496
DB 181 ACTGTGGGCAATGGCAGTTGGTATATCTGTACACACCTTCTGGTATTTGATATTTTA 240
QY 2497 GCACTTGTGTTATCCGCATPAAAGAGATPAAAGCAAGATATATGTTGAAGTCTCAA 2556
DB 241 GCACTTGTGTTATCCGCATPAAAGAGATPAAAGCAAGATATATGTTGAAGTCTCAA 300
QY 2557 GCATTCGAATCAAAACCTCTGAGAGAGTGAATTTGAAAAGATGTTGAATTTATATAG 2616
DB 301 GCATTCGAATCAAAACCTCTGAGAGAGTGAATTTGAAAAGATGTTGAATTTATATAG 360
QY 2617 CAAGTGTATTTGAGCAACACATCTCATCTTATCTTATCTTATCTTATCTTATCTTAT 2676
DB 361 CAAGTGTATTTGAGCAACACATCTCATCTTATCTTATCTTATCTTATCTTATCTTAT 420
QY 2677 ATTTTAAACAGATATTCCTCTGCTTATATATTTGCTAATATTTCTTTTGTAG 2736
DB 421 ATTTTAAACAGATATTCCTCTGCTTATATATTTGCTAATATTTCTTTTGTAG 480
QY 2737 GTGAGTCTTGCTCTGCGCCAGCGTGAGTACAGTG 2775
DB 481 GTGAGTCTTGCTCTGCGCCAGCGTGAGTACAGTG 519

RESULT 10
AAA16509
ID AAA16509 standard; DNA; 627 BP.
XX
AC AAA16509;
XX
DT 14-JUN-2000 (first entry)
XX
DE Human colon cancer differentially expressed nucleotide sequence #514.
XX
KW Colon cancer; detect; differential expression; human; treatment;
KM detect mutation; non-invasive diagnostic method; ds.
XX
OS Homo sapiens.
XX
XX WO200012702-A2.
XX
XX 09-MAR-2000.
XX
XX 30-AUG-1999; 99WO-US19424.
XX
XX 31-AUG-1998; 98US-0098639.
XX
XX 27-JAN-1999; 99US-0117393.
XX
XX (FARB) BAYER CORP.
XX
PI Endege WO, Steinmann KE, Aetle JH, Burgess CC, Carroll E,
PI Cactino TJ, Dwivedi P, Ford DM, Lewis ME, Molino GA, Monahan JE,
PI Schlegel R;

XX WPI; 2000-256641/22.
XX

PT Novel nucleic acid and proteins for identifying therapeutic agents
XX useful for treating and diagnosing cancer, especially colon cancer -
XX
XX
PS Claim 16; Page 332; 345pp; English.

CC This sequence represents a human nucleotide sequence which is
CC differentially expressed in colon cancer cells compared to the expression
CC levels in normal cells. The nucleotide sequence can be used as a source
CC of primers and probes. The nucleotide sequence is useful for determining
CC the phenotype of a cell by detecting the differential expression of the
CC sequence relative to a normal cell. The probes derived from the sequence
CC can also be used to determine the phenotype of cells in a sample. Probes
CC and antibodies which hybridise to the nucleotide sequence can also be
CC used to determine the phenotype of a cell. The primers are useful for
CC detecting a mutation in a test nucleotide sequence and also for detecting
CC cancer, preferably colon cancer. Antibodies against the protein encoded
CC by the nucleotide sequence can also be used in a method to detect colon
CC cancer. The diagnostic method is non-invasive and accurate for diagnosing
CC colon cancer at an early stage.

XX Sequence 627 BP; 159 A; 134 C; 146 G; 179 T; 9 other;
SQ

Query Match 15.3%; Score 511.4; DB 21; Length 627;
Best Local Similarity 99.0%; Pred. No. 2.8e-126;
Matches 512; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2254 GGTACTCATCCCGACTGTCTACACGACACAGACTTTGAGAGAGGGGTATGTCGTC 2313
DB 1 GGTACTCATCCCGACTGTCTACACGACACAGACTTTGAGAGAGGGGTATGTCGTC 60
QY 2314 TTGATCCGATCAATATATGAGGGGTCGCCACCTTGGAGGCAATGTTCTTTACCAAGTT 2373
DB 61 TTGATCCGATCAATATATGAGGGGTCGCCACCTTGGAGGCAATGTTCTTTACCAAGTT 120
QY 2374 ACATTCGAGTGTGTGGAAGAGTGTTCGCGCAGAGGTACACAGACTGGGATA 2433
DB 121 ACATTCGAGTGTGTGGAAGAGTGTTCGCGCAGAGGTACACAGACTGGGATA 180
QY 2434 CCCACTGTGGGCAATGGCAGTTGTATATCTGTACACACCTTCTGTGATTTGTATATTT 2493
DB 181 CCCACTGTGGGCAATGGCAGTTGTATATCTGTACACACCTTCTGTGATTTGTATATTT 240
QY 2494 TTGACAGTTGTGTTATCCGCATPAAAGAGATPAAAGCAAGATATGTTGAAGTCT 2553
DB 241 TTGACAGTTGTGTTATCCGCATPAAAGAGATPAAAGCAAGATATGTTGAAGTCT 300
QY 2554 CAAGCATCTGAGTCAAAACCTCTGAGAGCTGAATTTGAAAAGATGTTGAATTTATTA 2613
DB 301 CAAGCATCTGAGTCAAAACCTCTGAGAGCTGAATTTGAAAAGATGTTGAATTTATTA 360
QY 2614 TAGCAAGTCTATTTGAGCAACACATCTCATCTTATCTTATCTTATCTTATCTTATCT 2673
DB 361 TAGCAAGTCTATTTGAGCAACACATCTCATCTTATCTTATCTTATCTTATCTTATCT 420
QY 2674 ATTAATTTTAAACAGATATTCCTCTGCTTATATATTTGCTAATATTTCTTTT 2733
DB 421 ATTAATTTTAAACAGATATTCCTCTGCTTATATATTTGCTAATATTTCTTTT 480
QY 2734 GAGGTGGAGTCTTGCTCTGTCGCCAGAGCTGAGATAC 2770
DB 481 GAGGTGGAGTCTTGCTCTGTCGCCAGAGCTGAGATAC 517

RESULT 11
ABL38503
ID ABL38503 standard; cDNA; 501 BP.
XX
AC ABL38503;
XX
DT 08-APR-2002 (first entry)

XX Human colon tumour antigen polynucleotide SEQ ID NO:2092.
 DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 XX Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 XX Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 OS Homo sapiens.
 PN WO200196388-A2.
 PD 20-DEC-2001.
 XX 08-JUN-2001; 2001WO-US18557.
 PF 08-JUN-2001; 2000US-210899P.
 XX 09-JUN-2000; 2000US-210899P.
 PR 20-FEB-2001; 2001US-270216P.
 XX (CORI-) CORIXA CORP.
 PA Jiang Y, Harlocker SL, Secretist H;
 PI WPI; 2002-114514/15.
 XX Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX Claim 1; SEQ ID 2092; 105dp; English.
 PS ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
 CC which were isolated from human colon tumour and colon metastatic tumor
 CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
 CC production. (I) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (I) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (I)
 CC can be used in the diagnosis of a colon tumour.
 XX Sequence 501 BP; 130 A; 99 C; 110 G; 159 T; 3 other;
 SQ Query Match 14.9%; Score 497.2; DB 24; Length 501;
 Best Local Similarity 99.6%; Pred. No. 1.5e-122;
 Matches 497; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2257 ACTGATGCGGACGCTGTCTACAGGACACAGACTTTGAGAGAGGCGGTATGCTCTTG 2316
 DB 1 ACTGATGCGGACGCTGTCTACAGGACACAGACTTTGAGAGAGGCGGTATGCTCTTG 60
 QY 2317 ATCCGATCATATGATGGGGGCGGCGCCACCTTGGAGAGGCAATGTTCTTACCAAGTTACA 2376
 DB 61 ATCCGATCATATGATGGGGGCGGCGCCACCTTGGAGAGGCAATGTTCTTACCAAGTTACA 120
 QY 2377 TTCTGCAAGTTGTGGAAGAGTTGTTCCGGCAGAGAGGTCAACAGCTGGGATACC 2436
 DB 121 TTCTGCAAGTTGTGGAAGAGTTGTTCCGGCAGAGAGGTCAACAGCTGGGATACC 180
 QY 2437 ACTGATGCGGACGAGTGTATATCTGTCACACCTTCTGTGATTTGATTAATTTTA 2496
 DB 181 ACTGATGCGGACGAGTGTATATCTGTCACACCTTCTGTGATTTGATTAATTTTA 240
 QY 2497 GCAATGTTGTTATCCGATTAAGAAGATTAAGGCAAGATTAAGTGAAGTCTCAA 2556
 DB 241 GCAATGTTGTTATCCGATTAAGAAGATTAAGGCAAGATTAAGTGAAGTCTCAA 300
 QY 2557 GCATCTGAAGTCAACCTCTGAGAGTGAATTTGAAGATTTGATTAATTAATAG 2616
 DB 301 GCATCTGAAGTCAACCTCTGAGAGTGAATTTGAAGATTTGATTAATTAATAG 360
 QY 2617 CAAAGTCTATTGAGCAACCAATCTCATCTTATTTCACTTAAGCTGATTA 2676
 DB 361 CAAAGTCTATTGAGCAACCAATCTCATCTTATTTCACTTAAGCTGATTA 420

QY 2677 ATTTTAAACAGATATCCCTCTGCTCTTAATATTTGCAATATTTCTTTTGGAG 2736
 DB 421 ATTTTAAACAGATATCCCTCTGCTCTTAATATTTGCAATATTTCTTTTGGAN 480
 QY 2737 GTGGAGTCTTGTCTGTGCG 2755
 DB 481 GTGGAGTCTTGTCTGTGCG 499
 RESULT 12
 ABL36982/c
 ID ABL36982 standard; cDNA; 495 BP.
 XX ABL36982;
 AC ABL36982;
 DT 08-APR-2002 (first entry)
 XX Human colon tumour antigen polynucleotide SEQ ID NO:571.
 DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 XX Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
 OS Homo sapiens.
 PN WO200196388-A2.
 PD 20-DEC-2001.
 XX 08-JUN-2001; 2001WO-US18557.
 PF 08-JUN-2001; 2000US-210899P.
 PR 20-FEB-2001; 2001US-270216P.
 XX (CORI-) CORIXA CORP.
 PA Jiang Y, Harlocker SL, Secretist H;
 PI WPI; 2002-114514/15.
 XX Novel isolated colon tumor polynucleotide differentially expressed in
 PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
 PT useful for inhibiting development of cancer in patient -
 XX Claim 1; SEQ ID 571; 105dp; English.
 PS ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
 CC which were isolated from human colon tumour and colon metastatic tumor
 CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
 CC production. (I) can be used for stimulating and/or expanding T cells
 CC specific for a tumour protein on contact with the T cells. They are also
 CC useful for inhibiting the development of cancer in a patient. (I) can be
 CC used as probes or primers for nucleic acid hybridisation, for preparing
 CC mutant species primers, or primers for use in genetic constructions. (I)
 CC can be used in the diagnosis of a colon tumour.
 XX Sequence 495 BP; 156 A; 114 C; 93 G; 126 T; 6 other;
 SQ Query Match 14.0%; Score 469.4; DB 24; Length 495;
 Best Local Similarity 98.0%; Pred. No. 4.1e-115;
 Matches 483; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
 QY 2278 AGGACACAGACTTTGAGAGAGGCGGTATGCTCTTATCCGATCAATGATGGGGT 2337
 DB 493 AGGACACAGACTTTGAGAGAGGCGGTATGCTCTTATCCGATCAATGATGGGGT 434
 QY 2338 CGGCAACCTTGAAGGCAATGTTCTTACCAAGTATCTGCAAGTTGTGGAAGA 2397
 DB 433 CGGCAACCTTGAAGGCAATGTTCTTACCAAGTATCTGCAAGTTGTGGAAGA 374
 QY 2398 AGTTGTTCCGGCAGAGGTCAACAGAC--TGGATATCCCATCTGTGGGATGGCAATTG 2455
 DB 373 AGTTGTTCCGGCAGAGGTCAACAGANCTGGGATTAACCATCTGTGGGATGGCAATTG 314

QY 2456 GTATATGCTGACCAACCTCTGTGATGGTATATTTAGCAGTTGTGTTATCCGCA 2515
DB 313 GTATATGCTGACCAACCTCTGTGATGGTATATTTAGCAGTTGTGTTATCCGCA 254
QY 2516 TTAAGAAGATTAAGGCAAAATATATTTGAAAGTGTCAAGCATCTGAAGTCAAACTC 2575
DB 253 TTAAGAAGATTAAGGCAAAATATATGTTGAAAGTGTCAAGCATCTGAAGTCAAACTC 194
QY 2576 TGAGAAGCTGAATTTGAAAAGAAATGTTGAATTTATAGCAAGTCTATTTGAGCAAC 2635
DB 193 TGAGAAGCTGAATTTGAAAAGAAATGTTGAATTTATAGCAAGTCTATTTGAGCAAC 134
QY 2636 AACCATCTCATCTATTACTTTTTCATCTAACGTGCATTAATATTTTAAACAGATAATC 2695
DB 133 AACCATCTCATCTATTACTTTTTCATCTAACGTGCATTAATATTTTAAACAGATAATC 74
QY 2696 CCTTTGTCTTTTAATATTTTCTAAATTTTCTTTTGTAGGTGAGTCTGTCTGTG 2755
DB 73 CCTTTGTCTTTTAATATTTTCTAAATATTTTCTTTTGTAGGTGAGTCTGTCTGTG 14
QY 2756 CCCAGGCTGAGT 2768
DB 13 CCCAGGCTGAGT 1
RESULT 13
ABL36509
ID ABL36509 standard; cDNA; 492 BP.
XX ABL36509;
AC
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:98.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX colon tumour metastatic antigen; diagnosis; gene; ss.
OS Homo sapiens.
XX
XX NO200196388-A2.
PN
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
XX
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
DR WPI; 2002-114514/15.
XX
XX
PT Novel isolated colon tumor polynucleotide differentially expressed in
XX colon tumor or colon metastatic tumor and polypeptides encoded by them,
XX useful for inhibiting development of cancer in patient -
XX
PS Claim 1; SEQ ID 98; 105bp; English.
XX
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
XX which were isolated from human colon tumour and colon metastatic tumor
XX cDNA libraries. (I) have cytostatic activity and can be used in vaccine
XX production. (I) can be used for stimulating and/or expanding T cells
XX specific for a tumour protein on contact with the T cells. They are also
XX useful for inhibiting the development of cancer in a patient. (I) can be
XX used as probes or primers for nucleic acid hybridisation, for preparing
XX mutant species primers, or primers for use in genetic constructions. (I)
XX can be used in the diagnosis of a colon tumour.
XX
SQ Sequence 492 BP; 132 A; 94 C; 109 G; 155 T; 2 other;

Query Match 13.8%; Score 461.6; DB 24; Length 492;
Best Local Similarity 99.0%; Pred. No. 5e-113;
Matches 485; Conservative 1; Mismatches 1; Indels 3; Gaps 2;
QY 2257 ACTCATCCCGACTGTCTTACCAAGCAGACAGACTTTGAGAGAGGGCGTATGCTTTG 2316
DB 1 ACTCATCCCGACTGTCTTACCAAGCAGACAGACTTTGAGAGAGGGCGTATGCTTTG 60
QY 2217 ATCCGATCAATGATGGGGGTGGGCCACCCCTTGAAGGGCATTTGTTTACCAAGTTACA 2376
DB 61 ATCCGATCAATGATGGGGGTGGGCCACCCCTTGAAGGGCATTTGTTTACCAAGTTACA 120
QY 2377 TTCTGCAATTTGTGGAAGAAATTTTCCGCGCAGACAGTCAACAGACTGGGATACC 2436
DB 121 TTCTGCAATTTGTGGAAGAAATTTTCCGCGCAGACAGTCAACAGACTGGGATACC 180
QY 2437 ACTGTGGGCAATGGCAGTTGGTATCTGCTGACCAACCTTGTGATTTGATTAATTTA 2496
DB 181 ACTGTGGGCAATGGCAGTTGGTATCTGCTGACCAACCTTGTGATTTGATTAATTTA 240
QY 2497 GCAGTTGTGTTATCCGATTAAGAAAGATTAAGGCAAAAGATTAATGTTGAAGTGCCTCA 2556
DB 241 GCAGTTGTGTTATCCGATTAAGAAAGATTAAGGCAAAAGATTAATGTTGAAGTGCCTCA 300
QY 2557 GCATCTGAAGTCAAACTCTGAGAAAGCTGAATTTGAAGAAATTTGAATTTATATAG 2616
DB 301 GCATCTGAAGTCAAACTCTGAGAAAGCTGAATTTGAAGAAATTTGAATTTATATAG 360
QY 2617 CAAGTGTATTTACGACAAACCATCTCATCTTAATTTCACTTTCACTAACGTGCATTA 2676
DB 361 CAAGTGTATTTACGACAAACCATCTCATCTTAATTTCACTTTCACTAACGTGCATTA 420
QY 2677 A--TTTTTAAAGAAATTTCCCTCTGTGCC--TTTAAATATTCCTAAATTTCTTTT 2733
DB 421 AATTTTAAAGAAATTTCCCTCTGTGCC--TTTAAATATTCCTAAATTTCTTTT 480
QY 2734 GAGGTGAGT 2743
DB 481 GAGGTGAGT 490
RESULT 14
ABK44858
ID ABK44858 standard; cDNA; 516 BP.
XX
XX ABK44858;
AC
XX
DT 05-JUN-2002 (first entry)
XX
XX
DE cDNA encoding colon tumour protein, SEQ ID NO 409.
XX
XX
KW Human; colon tumour; vaccine; colon cancer; immunogenic;
XX immunotherapy; gene; ss.
OS Homo sapiens.
XX
XX
PN WO200212328-A2.
XX
PD 14-FEB-2002.
XX
PF 31-JUL-2001; 2001WO-US24218.
XX
PR 03-AUG-2000; 2000US-223283P.
XX
PR 28-MAR-2001; 2001US-279763P.
XX
PR 29-JUN-2001; 2001US-302051P.
XX
PA (CORI-) CORIXA CORP.
XX
PI King GE, Meagher MJ, Xu J, Secrist H;
XX
DR WPI; 2002-241739/29.
XX
XX
PT New colon cancer polypeptides and polynucleotides, useful as vaccines,

PT for diagnosing, preventing, and treating colon cancer, and as markers
for the progression of cancer -

XX Claim 1; SEQ ID No 409; 147bp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
compositions, such as vaccines, for the diagnosis, prevention, and
treatment of colon cancer. Polynucleotide sequences may be used as
hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
proteins in tumour cells. The compositions are useful for stimulating an
immune response against cancer, particularly for the immunotherapy of
colon cancer, and as markers for the progression of cancer.

CC ABR4450-ABK46237 represent coding sequences of human colon tumour
proteins of the invention.

CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
for this patent did not form part of the printed specification but was
supplied by the European Patent Office.

XX Sequence 516 BP; 166 A; 112 C; 113 G; 125 T; 0 other;

Query Match 12.9%; Score 431.4; DB 24; Length 516;
Best Local Similarity 99.8%; Pred. No. 6e-105;
Matches 432; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 570 CCAGCTTTATTCAGATGTCATCCAGCTTCCCATGATCAACATGTCATCTTCA 629
DB 1 CCAGCTTTATTCAGATGTCATCCAGCTTCCCATGATCAACATGTCATCTTCA 60

QY 630 GATCAACACAAACGGAGGAGCATCTCTTACCCGAGAGGATCTCAGAAATGAATCC 689
DB 61 GATCAACACAAACGGAGGAGCATCTCTTACCCGAGAGGATCTCAGAAATGAATCC 120

QY 690 TGCTAAGATCTCTCTATATCTGTGATCTCAGTGAAGACATGGAGCCAGAGTGA 749
DB 121 TGCTAAGATCTCTCTATATCTGTGATCTCAGTGAAGACATGGAGCCAGAGTGA 180

QY 750 GAATTCCTTCAGTATACCAATCTGTGATATCAATAGTGAAGAAATATTTGAAAGC 809
DB 181 GAATTCCTTCAGTATACCAATCTGTGATATCAATAGTGAAGAAATATTTGAAAGC 240

QY 810 ACCAAACCTGTGAGATGTGAGAAACTCACTGATCTCAACCCATCAAAATCACTCA 869
DB 241 ACCAAACCTGTGAGATGTGAGAAACTCACTGATCTCAACCCATCAAAATCACTCA 300

QY 870 GGTGCGGTGAATGATCCCGGTGACCAATATTTCTTGTAGTGAAGAAAGCTGCCAAG 929
DB 301 GGTGCGGTGAATGATCCCGGTGACCAATATTTCTTGTAGTGAAGAAAGCTGCCAAG 360

QY 930 ATTCCATTTTCATATGACCAAGAGATTTTAACTGATCTGAGCCCTTGGACCGAGA 989
DB 361 ATTCCATTTTCATATGACCAAGAGATTTTAACTGATCTGAGCCCTTGGACCGAGA 420

QY 990 AGAAAGATGCA 1002
DB 421 AGAAAGATGCA 433

RESULT 15
ABK45154
ID ABR45154 standard; cDNA; 568 BP.

XX ABR45154;

XX 05-JUN-2002 (first entry)

DE cDNA encoding colon tumour protein, SEQ ID No 705.

KW Human; colon tumour; vaccine; colon cancer; immunogenic;

KW immunotherapy; gene; ss.

XX Homo sapiens.

XX PN W020021328-A2.

XX PD 14-FEB-2002.

XX PF 31-JUL-2001; 2001WO-US24218.

XX PR 03-AUG-2000; 2000US-223283P.

XX PR 28-MAR-2001; 2001US-279763P.

XX PR 29-JUN-2001; 2001US-302051P.

XX PA (CORI-) CORIXA CORP.

XX PI King GE, Meagher MJ, Xu J, Secretist H;

XX PT New colon cancer polypeptides and polynucleotides, useful as vaccines,
for diagnosing, preventing, and treating colon cancer, and as markers
for the progression of cancer -

XX PS Claim 1; SEQ ID No 705; 147bp; English.

XX The invention relates to polynucleotides encoding colon tumour proteins.
CC The polynucleotides and encoded polypeptides are useful in pharmaceutical
compositions, such as vaccines, for the diagnosis, prevention, and
treatment of colon cancer. Polynucleotide sequences may be used as
hybridisation probes or primers, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of tumour polypeptides and
proteins in tumour cells. The compositions are useful for stimulating an
immune response against cancer, particularly for the immunotherapy of
colon cancer, and as markers for the progression of cancer.

CC ABR4450-ABK46237 represent coding sequences of human colon tumour
proteins of the invention.

CC Note: With the exception of SEQ ID No 1 and 2, the sequence data
for this patent did not form part of the printed specification but was
supplied by the European Patent Office.

XX Sequence 568 BP; 159 A; 125 C; 94 G; 190 T; 0 other;

Query Match 12.3%; Score 410.4; DB 24; Length 568;
Best Local Similarity 99.8%; Pred. No. 2.6e-99;
Matches 411; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2922 CCAGGCTGCTTGAATCTCTGACGTCAAGTATGCTCCGCTGCTCCCAATPACAG 2981
DB 1 CCAGGCTGCTTGAATCTCTGACGTCAAGTATGCTCCGCTGCTCCCAATPACAG 60

QY 2982 GCATGAACCACTGACCCCACTTATGATATTTCAATGTCTATAGACATTAAGAGATT 3041
DB 61 GCATGAACCACTGACCCCACTTATGATATTTCAATGTCTATAGACATTAAGAGATT 120

QY 3042 TTTCAATTTTCATGATATTTTCTCTCTGCAATGCGTTAGTACTTGTGTTTTCC 3101
DB 121 TTTCAATTTTCATGATATTTTCTCTCTGCAATGCGTTAGTACTTGTGTTTTCC 180

QY 3102 TTTTGGGGCAAGACAGATCAATTAATATCTGATCAATTTTCTTATACAGAGATAT 3161
DB 181 TTTTGGGGCAAGACAGATCAATTAATATCTGATCAATTTTCTTATACAGAGATAT 240

QY 3162 ATCAGTGTGTCTCATGAACTGCTGATTCATTTATGTTTTTCTGATTCATCTG 3221
DB 241 ATCAGTGTGTCTCATGAACTGCTGATTCATTTATGTTTTTCTGATTCATCTG 300

QY 3222 TGTCCCTTCATCTTGAATCTCTTGGTATTTTCACTGAATTTCAACATTTGTCAAGAA 3281
DB 301 TGTCCCTTCATCTTGAATCTCTTGGTATTTTCACTGAATTTCAACATTTGTCAAGAA 360

QY 3282 GAAAAAAGTGAAGATCTCAGAAAAATTAATTAATTAAGAACGCTTTTGC 3333
DB 361 GAAAAAAGTGAAGATCTCAGAAAAATTAATTAATTAAGAACGCTTTTGC 412

Search completed: June 21, 2003, 21:24:22
Job time : 705 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 21:10:19 ; Search time 183 Seconds
(without alignments)
5605.649 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345
Sequence: 1 ggaatccgcgcgcgcacacg.....cctttcgccgcgcgaatcc 3345

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2499	74.7	2499	1	US-08-431-560-2
2	2499	74.7	2499	1	US-08-463-345-2
3	511.4	15.3	627	4	US-09-385-982-514
4	331.4	9.9	346	4	US-09-221-298-21
5	284.8	8.5	591	4	US-09-385-982-300
6	191.8	5.7	36741	4	US-09-301-665-3
7	191.2	5.7	162450	4	US-09-345-882-1
8	191	5.7	4853	4	US-08-881-450A-22
9	190.8	5.7	14581	4	US-08-520-373D-4
10	190.8	5.7	22481	4	US-08-367-841A-43
11	190.8	5.7	22481	5	PCT-US95-07201-43
12	190.8	5.7	22484	4	US-09-875-223-2
13	190.6	5.7	98844	4	US-09-791-211-10
14	190.4	5.7	8453	4	US-09-167-881-45
15	190.4	5.7	36651	4	US-09-738-894A-3
16	189.6	5.7	21234	4	US-09-810-671-3
17	188.6	5.6	70000	4	US-09-851-896-3
18	188	5.6	42571	4	US-09-810-347-3
19	188	5.6	111282	4	US-09-754-250-3
20	186.6	5.6	62804	4	US-09-800-960-3
21	186.4	5.6	956	4	US-09-641-638-56
22	186.4	5.6	20674	4	US-09-641-638-651
23	186.4	5.6	31571	1	US-08-323-443B-1
24	186.4	5.6	53526	3	US-08-658-136-2
25	186.4	5.6	53577	3	US-08-658-136-1
26	186.4	5.6	62804	4	US-09-800-960-3
27	186	5.6	3627	4	US-09-323-873A-6

28	185.6	5.5	631	4	US-09-385-982-354	Sequence 354, App
29	185.4	5.5	87350	3	US-08-781-891-79	Sequence 79, App1
30	185.4	5.5	87543	4	US-09-791-211-3	Sequence 3, App1
31	185	5.5	7676	1	US-08-451-777A-7	Sequence 7, App1
32	185	5.5	7676	2	US-08-451-778A-7	Sequence 7, App1
33	185	5.5	7676	5	PCT-US95-06743-7	Sequence 7, App1
34	185	5.5	7676	5	US-08-998-208-7	Sequence 7, App1
35	184.8	5.5	36741	4	US-09-301-665-3	Sequence 3, App1
36	184.4	5.5	12047	2	US-09-022-461-1	Sequence 3, App1
37	184.4	5.5	12047	4	US-09-033-556-3	Sequence 3, App1
38	184.4	5.5	99500	4	US-09-798-096-10	Sequence 10, App1
39	184.2	5.5	43950	4	US-09-735-934A-3	Sequence 3, App1
40	184.2	5.5	55827	4	US-09-813-133A-3	Sequence 3, App1
41	183.6	5.5	1001	4	US-09-641-638-406	Sequence 406, App
42	183.2	5.5	72604	4	US-09-268-992-7	Sequence 7, App1
43	183.2	5.5	72604	4	US-09-657-474-7	Sequence 7, App1
44	182.4	5.5	6235	4	US-09-305-384-5	Sequence 5, App1
45	182.4	5.5	6679	4	US-09-305-384-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-08-431-560-2
; Sequence 2, Application US/08411560
; Patent No. 5620855
; GENERAL INFORMATION:
; APPLICANT: Anne H. Dantzig, et al.
; TITLE OF INVENTION: Mammalian Intiflux Peptide
; NUMBER OF INVENTIONS: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,560
; FILING DATE: 01-May-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,462
; FILING DATE: 04-FEB-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2499 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-431-560-2
Query Match 74.7%; Score 2499; DB 1; Length 2499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2499; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 ATGATCTTACGAGCCATCTTCACTCCCTGTCTCTTTATGCTTTATTTGGCAACTGGA 147
1 ATGATCTTACGAGCCATCTTCACTCCCTGTCTCTTTATGCTTTATTTGGCAACTGGA 60
QY 148 TATGGCCAAAGGGGAGGTTAGTGACCCCGAAGCCAGACGATTTCTATTTATGAA 207
61 TATGGCCAAAGGGGAGGTTAGTGACCCCGAAGCCAGACGATTTCTATTTATGAA 120
DB 208 GGCCAAAGAACGAGTCAATATATTTCAAGTTAAGCCAAATCCCTGTGTGACTTTT 267

121 GGGCAAGAACCGATCAATTAATTCAGTTTAAGGCCAATCTCTGCTGACTTTT 180
1268 GAACTAATCTGGGAGACAGACAATATTTGTGATAGAACGGAGGACTTGTATTAAC 327
181 GAACTAATCTGGGAGACAGACAATATTTGTGATAGAACGGAGGACTTGTATTAAC 240
328 AACAGACCTTGGACAGAGGAAACAAGATCTAATCTCAATCTCCAGGTTGACAGCTTGAC 387
241 AACAGACCTTGGACAGAGGAAACAAGATCTAATCTCAATCTCCAGGTTGACAGCTTGAC 300
388 GCTAATGGAATTAATAGTGAAGGGTCCAGTCCCTATCACTAGAGTGAAGAGCATCAAC 447
301 GCTAATGGAATTAATAGTGAAGGGTCCAGTCCCTATCACTAGAGTGAAGAGCATCAAC 360
448 GACAAATGACCCACGTTCTTCAGTCAAGTACGAAAGCTCAGTAAGGCAAACTCTGCG 507
361 GACAAATGACCCACGTTCTTCAGTCAAGTACGAAAGCTCAGTAAGGCAAACTCTGCG 420
508 CCAAGAAAGCCCTTCTGTATGTCAATGCCACAGACTGATGATCCGGCCACTCCCAAT 567
421 CCAAGAAAGCCCTTCTGTATGTCAATGCCACAGACTGATGATCCGGCCACTCCCAAT 480
568 GGGCAGCTTATTAACAGATTTGCAATCAGCTTCCAGTCAAGATGATGATGATGATGAT 627
481 GGGCAGCTTATTAACAGATTTGCAATCAGCTTCCAGTCAAGATGATGATGATGATGAT 540
628 CAGATCAACAACAACGAGGAGCCATCTCTTAACCCGAGAGGAGCTCAGAGAAATTAAT 687
541 CAGATCAACAACAACGAGGAGCCATCTCTTAACCCGAGAGGAGCTCAGAGAAATTAAT 600
688 CTTCTCAAGATCTCTCTATATCTGTGATCTCAAGTGAAGGACATGGGAGGCAAGAT 747
601 CTTCTCAAGATCTCTCTATATCTGTGATCTCAAGTGAAGGACATGGGAGGCAAGAT 660
748 GAGAAATCTCTCAAGTGAATCACTGTGATCTCAAGTGAATCACTGTGATCTCAAGT 807
661 GAGAAATCTCTCAAGTGAATCACTGTGATCTCAAGTGAATCACTGTGATCTCAAGT 720
808 GCAACAAAACCTGTGAGATGATGAGAAAATCAACTGATCTCTCAACCCCATCAAAATCACT 867
721 GCAACAAAACCTGTGAGATGATGAGAAAATCAACTGATCTCTCAACCCCATCAAAATCACT 780
868 CAGGTGGGTGGAATGATCCCGGTGACAAATATCTTAAATTTGACAAAGAACTGCA 927
781 CAGGTGGGTGGAATGATCCCGGTGACAAATATCTTAAATTTGACAAAGAACTGCA 840
928 AGATTTCCATTTCAATTTGACAGAGAGGAGATATTTAAGTCACTCAGCCCTTGACCGA 987
841 AGATTTCCATTTCAATTTGACAGAGAGGAGATATTTAAGTCACTCAGCCCTTGACCGA 900
988 GAAAGAAAAGATGATATGTTTTTATGACAGTGAAGATGATGATGATGATGATGATGAT 1047
901 GAAAGAAAAGATGATATGTTTTTATGACAGTGAAGATGATGATGATGATGATGATGAT 960
1048 TCATATCCGCTGGAATTCATGTAAGTAAAGTAAATTAATTAATTAATTAATTAATTAAT 1107
961 TCATATCCGCTGGAATTCATGTAAGTAAAGTAAATTAATTAATTAATTAATTAATTAAT 1020
1108 CCGTCACAGTAACCGTATTTGAGTCCAGAGATGAAGACATGGGGTAAACATGATGCGG 1167
1021 CCGTCACAGTAACCGTATTTGAGTCCAGAGATGAAGACATGGGGTAAACATGATGCGG 1080
1168 ACCCTTAATGACATGACAGGAGATGAAGAAAATCTGCAACAGTTTTCTAACTAAGG 1227
1081 ACCCTTAATGACATGACAGGAGATGAAGAAAATCTGCAACAGTTTTCTAACTAAGG 1140
1228 ATTGTGAGCAAACTCCCAACTTCCATGATGATGATGATGATGATGATGATGATGAT 1287
1141 ATTGTGAGCAAACTCCCAACTTCCATGATGATGATGATGATGATGATGATGATGAT 1200
1288 GGAATGTTACAGTTAGCTAAACAGTCTTGAAGAAAGCAAGATCTCCCTCAATCAACTTA 1347

1201 GGAATGTTACAGTTAGCTAAACAGTCTTGAAGAAAGCAAGATATCTCTCAATCAACTTA 1260
1348 AGCATAGAGTGTCTGACAAAGATTTCAAGACCCCTTTGTTTGTGAATTCAGACTTAT 1407
1261 AGCATAGAGTGTCTGACAAAGATTTCAAGACCCCTTTGTTTGTGAATTCAGACTTAT 1320
1408 GATATCAATGATCAGATATCCCATCTTTGAAAAATCAGATTTATGAAACCTGACTCTTGCT 1467
1321 GATATCAATGATCAGATATCCCATCTTTGAAAAATCAGATTTATGAAACCTGACTCTTGCT 1380
1468 GAAACACAAAATTTGGTTCACCAATCTTAACCATCAAGGCCATGATGCTGATGAGCCA 1527
1381 GAAACACAAAATTTGGTTCACCAATCTTAACCATCAAGGCCATGATGCTGATGAGCCA 1440
1528 TTTACTGGGAGTTCTAAATTTCTGATTCATATCAATAAGGAGACGTAGGGAGCCCTG 1587
1441 TTTACTGGGAGTTCTAAATTTCTGATTCATATCAATAAGGAGACGTAGGGAGCCCTG 1500
1588 GGGGTGACACAGATCCCATACCAACCCGATATGTCAATATTAATAAGCCTCTTGAT 1647
1501 GGGGTGACACAGATCCCATACCAACCCGATATGTCAATATTAATAAGCCTCTTGAT 1560
1648 TTTGAAACAGCAGCTGTTTCCAAATGTGTTCAAAACAGAAAATCTGAGCCTTAAGT 1707
1561 TTTGAAACAGCAGCTGTTTCCAAATGTGTTCAAAACAGAAAATCTGAGCCTTAAGT 1620
1708 TTTGTGTGAAGTACAAATGCAATGTTTTCGCAATCTTCAACGTTATGTGACAGATGTG 1767
1621 TTTGTGTGAAGTACAAATGCAATGTTTTCGCAATCTTCAACGTTATGTGACAGATGTG 1680
1768 AATGAAGACCTCAATTTTCCCAACAGTATTCGAAGCAAAAGTCACTGAGATGATAGCT 1827
1681 AATGAAGACCTCAATTTTCCCAACAGTATTCGAAGCAAAAGTCACTGAGATGATAGCT 1740
1828 ATAGGCACTAAAGTGGGCAATGTGATGCTGCCAAGATTCAGAAAGTCTGACATTAAGCTAT 1887
1741 ATAGGCACTAAAGTGGGCAATGTGATGCTGCCAAGATTCAGAAAGTCTGACATTAAGCTAT 1800
1888 TCACTGAGGGGAGACACAAAGAGTTGGCTTAAATTTGACACAGTGAATGATGATGAT 1947
1801 TCACTGAGGGGAGACACAAAGAGTTGGCTTAAATTTGACACAGTGAATGATGATGAT 1860
1948 AGTGTGGCTCAATTTGACAGAGAACCGGAAGTCCATATCGGATGACAAAGTGTGGCACA 2007
1861 AGTGTGGCTCAATTTGACAGAGAACCGGAAGTCCATATCGGATGACAAAGTGTGGCACA 1920
2008 GAAATGAGGGGAGTCTTCTTAAGCTGTGTGACAGTTCACCTGATCTTAATGATGTG 2067
1921 GAAATGAGGGGAGTCTTCTTAAGCTGTGTGACAGTTCACCTGATCTTAATGATGTG 1980
2068 AATGACAAACCTTCCAGGCTAGGCAAGACTACACGGGCTGTCTTCCGCAATCCCTC 2127
1981 AATGACAAACCTTCCAGGCTAGGCAAGACTACACGGGCTGTCTTCCGCAATCCCTC 2040
2128 AGTGAACCTGGAGTCAATTTGAGGCTAGTATGATGACACTAATTTGGGGT 2187
2041 AGTGAACCTGGAGTCAATTTGAGGCTAGTATGATGACACTAATTTGGGGT 2100
2188 CCCCATTTTACATTTTCCCTCGGACGTGAAGCTTACAAAAGCACTGGGAAGTTTCCAAA 2247
2101 CCCCATTTTACATTTTCCCTCGGACGTGAAGCTTACAAAAGCACTGGGAAGTTTCCAAA 2160
2248 ATCAATGATCTATGATGCGGACCTGTCTACAGGACACACAGCTTTAGAGAGGGGTAT 2307
2161 ATCAATGATCTATGATGCGGACCTGTCTACAGGACACACAGCTTTAGAGAGGGGTAT 2220
2308 GTGTCTTGAATCCGATCAATGATGAGGGGTGCGGCAACCTTGAAGGCACTTGTCTTTA 2367
2221 GTGTCTTGAATCCGATCAATGATGAGGGGTGCGGCAACCTTGAAGGCACTTGTCTTTA 2280
2368 CCAAGTTAATCTGCAATGATGATGAGAGAGATGTTTCCGACAGAGGTCAACAGACT 2427
2281 CCAAGTTAATCTGCAATGATGATGAGAGAGATGTTTCCGACAGAGGTCAACAGACT 2340

Oy	2428	GGGATACCACTGGTGGGCGATGGCATGTGTACTCTGCGACCACCTTCGTGTATTGCT	2487
Db	2341	GGGATACCACTGGGCGATGGCGATTGGTTACTCTGCGACCACTTCGTGTATTGCT	2400
Oy	2488	ATAATTTAGCAGTGTGTATTATCCGCTAAAGAAGTAAGAACAAGATAAAGTTGAA	2547
Db	2401	ATAATTTTAGCAGTGTGTATTATCCGCTAAAGAAGGTAAGAACAAGATAAAGTTGAA	2460
Oy	2548	AGTGCTCAAGCATCTGAAGTCAAACTCTGGAAGCTGA	2586
Db	2461	AGTGCTCAAGCATCTGAAGTCAAACTCTGGAAGCTGA	2499

```

RESULT 2
US-08-463-345-2
; Sequence 2, Application US/08463345
; Patent No. 5710018
; GENERAL INFORMATION:
; APPLICANT: Anne H. Dantzig, et al.
; TITLE OF INVENTION: Mammalian Influx Peptide
; TITLE OF INVENTION: Transporter
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,345
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,462
; FILING DATE: 04-FEB-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2499 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-463-345-2

```

Query Match	74.7%	Score 2499	DB 1	Length 2499
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2499	0	Mismatches 0	Indels 0	Gaps 0
Conservative				

Qy	88	ATGATATCTTCAGGGCCATCTTCATCCCTGGTCTCTTATAGCTTTATTTGGCAACTGGA	147
Db	1	ATGATATCTTCAGGGCCATCTTCATCCCTGGTCTCTTATAGCTTTATTTGGCAACTGGA	60
Qy	148	TATGGCGAAGAGGGGAAGTTTATGTGACCCCTGAAACCCATGACATTTTCTATTTATGAA	207
Db	61	TATGGCGAAGAGGGGAAGTTTATGTGACCCCTGAAACCCATGACATTTTCTATTTATGAA	120
Qy	208	GGCGCAAGAACGAGTCATTAATTATTCAGATTTAAGGCGATCTCTCGCTGAGACTTTT	267
Db	121	GGCGCAAGAACGAGTCATTAATTATTCAGATTTAAGGCGATCTCTCTGCTGAGACTTTT	180
Qy	268	GAACTPACTGGGGGAGCGAGCAACATATTTGTGATGAGACGGGAGGACCTTCGTATTC	327
Db	181	GAACTPACTGGGGGAGCGAGCAACATATTTGTGATGAGACGGGAGACTTCGTATTTAC	240
Qy	328	AAACAGCCTTGAGACGGGAAACAAGATCTACTACATCTCCAGGTGACGCTCTGGAC	387

D	241	AACAGAGCCTTGGACAGGGAAAACAAGATCTACTCAAAATCTCCAGGTGGACGCTTGAC	300
Q	388	GCTAATGAAATTATATGTGAGAGGCTCCAGTCCCTATCAACATAGAGTGAAGCATCAAC	447
D	301	GCTAAATGAATATATGTGAGAGGCTCCAGTCCCTATCAACATAGAGTGAAGCATCAAC	360
Q	448	GAACAATGACCAACGCTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGCGAACTCTGGC	507
D	361	GAACAATGACCAACGCTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGCGAACTCTGGC	420
Q	508	CCAGGAAAGCCCTTCTTGATGTCAATGSCACAACCTGGATGATCCGGCACTCCCAAT	567
D	421	CCAGGAAAGCCCTTCTTGATGTCAATGSCACAACCTGGATGATCCGGCACTCCCAAT	480
Q	568	GGCCAGCTTATTAACAGATTGTGATCCAGCTTCCATGATCAACAATGTCATGTACTTT	627
D	481	GGCCAGCTTATTAACAGATTGTGATCCAGCTTCCATGATCAACAATGTCATGTACTTT	540
Q	628	CAGATCAACAACAAAAGGAGGCCATCTCTTTACCCGAAGGGATCTCAGAAATTTGAAT	687
D	541	CAGATCAACAACAAAAGGAGGCCATCTCTTTACCCGAAGGGATCTCAGAAATTTGAAT	600
Q	688	CCGTGTAAGAAATCCTTCTATATCTGGTGTCTCAGTGAAGCAATGGAGGCGAGGT	747
D	601	CCGTGTAAGAAATCCTTCTATATCTGGTGTCTCAGTGAAGCAATGGAGGCGAGGT	660
Q	748	GAGAAATTCCTTCAGTGATACCAATCTGTGGATATCATGTGACAGAGAAATATTGGAAA	807
D	661	GAGAAATTCCTTCAGTGATACCAATCTGTGGATATCATGTGACAGAGAAATATTGGAAA	720
Q	808	GCAACCAAAACCTGTGAGATGTGGAAAACCTCAACGTATCTCAACCCCATCAAAATCACT	867
D	721	GCAACCAAAACCTGTGAGATGTGGAAAACCTCAACGTATCTCAACCCCATCAAAATCACT	780
Q	868	CAGGTGCGGTGAAAGATCCCGGTGACAAATATTTCTTATTTGAACAAAGAAAGCTGGCA	927
D	781	CAGGTGCGGTGAAAGATCCCGGTGACAAATATTTCTTATTTGAACAAAGAAAGCTGGCA	840
Q	928	AGATTCCCAATTTTCAATTGACAGAGAGAGATTTTACGTGACTGACCTTGGACGGA	987
D	841	AGATTCCCAATTTTCAATTGACAGAGAGAGATTTTACGTGACTGACCTTGGACGGA	900
Q	988	GAGAAAGAGATGATATGTTTTTATGCAAGTTGCAAGGATGAGTACGGAAAACCACTT	1047
D	901	GAGAAAGAGATGATATGTTTTTATGCAAGTTGCAAGGATGAGTACGGAAAACCACTT	960
Q	1048	TCATATCCGCTGGAATTCATGTAAAGTTAAAGATTTAAATATTAATATCACTACATGT	1107
D	961	TCATATCCGCTGGAATTCATGTAAAGTTAAAGATTTAAATATTAATATCACTACATGT	1020
Q	1108	CCGTCAACAGTAAACGATTTTGAAGTCCAGAGAAAGAAACGATCGGTAACAGTATCGGG	1167
D	1021	CCGTCAACAGTAAACGATTTTGAAGTCCAGAGAAAGAAACGATCGGTAACAGTATCGGG	1080
Q	1168	ACCCTTACTGCAATGACAGGGATGAAGAAAATATCTGCCAAGTTTTCTAAACTACAGG	1227
D	1081	ACCCTTACTGCAATGACAGGGATGAAGAAAATATCTGCCAAGTTTTCTAAACTACAGG	1144
Q	1228	ATTGTGAGCAAACTCCCAAACTTCCATGATGACCTTCTCTAATCCAAACTATAGCT	1287
D	1141	ATTGTGAGCAAACTCCCAAACTTCCATGATGACCTTCTCTAATCCAAACTATAGCT	1200
Q	1288	GGAAATGTAACGTTAAGCTTAAACGCTCTTGAAGACGAAGATCTCTCAGTACAACTTA	1347
D	1201	GGAAATGTAACGTTAAGCTTAAACGCTCTTGAAGACGAAGATCTCTCAGTACAACTTA	1266
Q	1348	ACGATAGAGGTGTCTGCAACAAGATTTCAAAACCTTTGTTTTGTGTGCAAACTCAAGTAT	1407
D	1261	ACGATAGAGGTGTCTGCAACAAGATTTCAAAACCTTTGTTTTGTGTGCAAACTCAAGTAT	1322
Q	1408	GATATCAATGATCAGATCCCATCTTTGAAAAATCAGATTAAGAAAACCTGACTTGTGT	1467
D	1321	GATATCAATGATCAGATCCCATCTTTGAAAAATCAGATTAAGAAAACCTGACTTGTGT	1380

QY 1468 GAAAGCAAAACATTGGGCTCCACCATTTTAACCATCCAGGCCACTGATGCTATGAGCCA 1527
 DB 1381 GAAAGCAAAACATTGGGCTCCACCATTTTAACCATCCAGGCCACTGATGCTATGAGCCA 1440
 QY 1528 TTACTGGGAGTCTAAATTTCTGTATCATATCATTAAGGAGAGAGAGAGAGAGAGAGAG 1587
 DB 1441 TTACTGGGAGTCTAAATTTCTGTATCATATCATTAAGGAGAGAGAGAGAGAGAGAGAG 1500
 QY 1588 GGGGTTGACACAGATCCCATACCAACACCGGATATGTCATTAATAAAAGCCTCTTGAT 1647
 DB 1501 GGGGTTGACACAGATCCCATACCAACACCGGATATGTCATTAATAAAAGCCTCTTGAT 1560
 QY 1648 TTGAAAACAGACAGTGTTCCTCAACATGTGTCTTAAGACAAAATCCTGAGCCTTAAGT 1707
 DB 1561 TTGAAAACAGACAGTGTTCCTCAACATGTGTCTTAAGACAAAATCCTGAGCCTTAAGT 1620
 QY 1708 TTGGTGTGAAGTACATGCAAGTCTTTTGCCAAAGTTCACGCTTATTTGACAGATGTG 1767
 DB 1621 TTGGTGTGAAGTACATGCAAGTCTTTTGCCAAAGTTCACGCTTATTTGACAGATGTG 1680
 QY 1768 AATGAAGCACTCAATTTTCCCAACACGTAATCCAAAGCAAAAGTCACTGAGATGCT 1827
 DB 1681 AATGAAGCACTCAATTTTCCCAACACGTAATCCAAAGCAAAAGTCACTGAGATGCT 1740
 QY 1828 ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAAAGATCTGGAATTAAGCTAT 1887
 DB 1741 ATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAAAGATCTGGAATTAAGCTAT 1800
 QY 1888 TCACTGAGGGGAGACACAAAGAGTTGGCTTAAATTAATGACACGTAAGTGTGAGATCTT 1947
 DB 1801 TCACTGAGGGGAGACACAAAGAGTTGGCTTAAATTAATGACACGTAAGTGTGAGATCTT 1860
 QY 1948 AGTGTGCTCCATTGACAGAGAAAGCCGGAAGTCCATATGCGGTGACAAAGTGTGAGCA 2007
 DB 1861 AGTGTGCTCCATTGACAGAGAAAGCCGGAAGTCCATATGCGGTGACAAAGTGTGAGCA 1920
 QY 2008 GAAATGAGGGGGGTCTTCTTAAAGTCTGTGTCAAGATTTCCACTGATCTTAAGAGATG 2067
 DB 1921 GAAATGAGGGGGGTCTTCTTAAAGTCTGTGTCAAGATTTCCACTGATCTTAAGAGATG 1980
 QY 2068 AATGACAACCTCCAGAGCTGACCAAGAGCTTACAGAGGAGCTTCTTCCGACATCCCTC 2127
 DB 1981 AATGACAACCTCCAGAGCTGACCAAGAGCTTACAGAGGAGCTTCTTCCGACATCCCTC 2040
 QY 2128 AGTGCACCTGGAAGTCTCATTTTTCGAGGCTACTGATGATGATCAGCACTTAATTTCCG 2187
 DB 2041 AGTGCACCTGGAAGTCTCATTTTTCGAGGCTACTGATGATGATCAGCACTTAATTTCCG 2100
 QY 2188 CCCCATTTTACATTTTCCCTCGGAGAGTGAAGCTTACAAAACGACCTGGAAAGTTTCCAA 2247
 DB 2101 CCCCATTTTACATTTTCCCTCGGAGAGTGAAGCTTACAAAACGACCTGGAAAGTTTCCAA 2160
 QY 2248 ATCAATGATCTCAATGCCCCGAGCTGTACACAGGACACAGACTTTGAGAGAGGCGCTAT 2307
 DB 2161 ATCAATGATCTCAATGCCCCGAGCTGTACACAGGACACAGACTTTGAGAGAGGCGCTAT 2220
 QY 2308 GTGCTCTTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAAGGCAATTTCTTTA 2367
 DB 2221 GTGCTCTTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAAGGCAATTTCTTTA 2280
 QY 2368 CCAATTCATCTCTGCAATGTGTGTGGAAGAGTGTTCCTGGGCAAGAGCTCAACAACAT 2427
 DB 2281 CCAATTCATCTCTGCAATGTGTGTGGAAGAGTGTTCCTGGGCAAGAGCTCAACAACAT 2340
 QY 2428 GGAATACCACTGTGGGAGTGGAGTGTGTATCTGCTGACCAACCTTCTGTATTTGT 2487
 DB 2341 GGAATACCACTGTGGGAGTGGAGTGTGTATCTGCTGACCAACCTTCTGTATTTGT 2400
 QY 2488 ATAAATTTAGCAGTGTGTATTCGACATAAAGAGATTAAGGCAAGATTAATTTGAA 2547
 DB 2401 ATAAATTTAGCAGTGTGTATTCGACATAAAGAGATTAAGGCAAGATTAATTTGAA 2460

QY 2548 AGTGCACAGCATCTGAAGTCAAACTCTGAGAGCTGA 2586
 DB 2461 AGTGCACAGCATCTGAAGTCAAACTCTGAGAGCTGA 2499
 RESULT 3
 US-09-385-982-514
 ; Sequence 514, Application US/09385982
 ; Patent No. 6262314
 ; GENERAL INFORMATION:
 ; APPLICANT: ENDEGE, WILSON O., ET AL.
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCNDA-260XX
 ; CURRENT FILING DATE: 1999-08-30
 ; EARLIER APPLICATION NUMBER: 09/328,111
 ; EARLIER FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: 60/117,393
 ; EARLIER FILING DATE: 1999-01-27
 ; EARLIER APPLICATION NUMBER: 60/098,639
 ; NUMBER OF SEQ ID NOS: 544
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 514
 ; LENGTH: 627
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(627)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-385-982-514
 Query Match 15.3%; Score 511.4; DB 4; Length 627;
 Best Local Similarity 99.0%; Pred. No. 1e-132;
 Matches 512; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 2254 GTTACTCATGCCCGGAGTGTCTACAGGACACAGACTTTGAGAGAGGGGATGTCGTC 2213
 DB 1 GTTACTCATGCCCGGAGTGTCTACAGGACACAGACTTTGAGAGAGGGGATGTCGTC 60
 QY 2314 TTGATCCGATCAATGATGAGGGGTCCGACACCTTGAAGGCAATTTCTTTACAGATT 2373
 DB 61 TTGATCCGATCAATGATGAGGGGTCCGACACCTTGAAGGCAATTTCTTTACAGATT 120
 QY 2374 ACATTTGCAATTTGTGTGAAGAGTGTTCCTGCGCCAGCAGAGTCAACGACTGGGATA 2433
 DB 121 ACATTTGCAATTTGTGTGAAGAGTGTTCCTGCGCCAGCAGAGTCAACGACTGGGATA 180
 QY 2434 CCACATGTGGGAGTATGAGTGTGATAGCTGACACACCTTGTGGATTTGATTAAT 2493
 DB 181 CCACATGTGGGAGTATGAGTGTGATAGCTGACACACCTTGTGGATTTGATTAAT 240
 QY 2494 TTAGCAGTGTGTTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCT 2553
 DB 241 TTAGCAGTGTGTTTATCCGATTAAGAGATTAAGGCAAGATTAATGTTGAAGTCT 300
 QY 2554 CAAGCATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAAGAAATTTGAATTTATA 2613
 DB 301 CAAGCATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAAGAAATTTGAATTTATA 360
 QY 2614 TAGCAAGTGTATTTAGCAAAACCATGTCATCACTTATTAATTTGATTAAGTGAAT 2673
 DB 361 TAGCAAGTGTATTTAGCAAAACCATGTCATCACTTATTAATTTGATTAAGTGAAT 420
 QY 2674 ATAAATTTTAAACAGATATTCCTCTGTGCTTAAATTTGCTAAATTTCTTTT 2733
 DB 421 ATAAATTTTAAACAGATATTCCTCTGTGCTTAAATTTGCTAAATTTCTTTT 480
 QY 2734 GAGGTGAGTCTGCTGTGTGCGCCAGGCTGAGATC 2770
 DB 481 GAGGTGAGTCTGCTGTGTGCGCCAGGCTGAGATC 517

```
RESULT 4
US-09-221-298-21/c
; Sequence 21, Application US/09221298
; Patent No. 6284241
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER
; FILE REFERENCE: 210121.471
; CURRENT APPLICATION NUMBER: US/09/221,298
; CURRENT FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (257)
; OTHER INFORMATION: where n is a, c, g or t
US-09-221-298-21

Query Match
Best Local Similarity 9.9%; Score 331.4; DB 4; Length 346;
Matches 343; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1661 CTGTTTCCACATTGTGTTCAGAGCAAAATCTGAGCCTTCTGTTGTGTGAGT 1720
DB 346 CTGTTTCCACATTGTGTTCAGAGCAAAATCTGAGCCTTCTGTTGTGTGAGT 287
QY 1721 ACAATGCAAGTCTTTTGGCAAGTTACG-CCTATTGTGACAGATGGAATGAAGCACT 1779
DB 286 ACAATGCAAGTCTTTTGGCAAGTTACG-CCTATTGTGACAGATGGAATGAAGCACT 227
QY 1780 CAATTTTCCACACGATTTTCCAGAGCAAAATGAGTGAAGATGAGTATGAGCACTAA 1839
DB 226 CAATTTTCCACACGATTTTCCAGAGCAAAATGAGTGAAGATGAGTATGAGCACTAA 167
QY 1840 GTGGGCAATGTACTGCGCAAGATCCAGAGGCTGTGACATTAAGTATTCAGTGAAGGGA 1899
DB 166 GTGGGCAATGTACTGCGCAAGATCCAGAGGCTGTGACATTAAGTATTCAGTGAAGGGA 107
QY 1900 GACACAGAGGTTGGTTAAATTGACAGTACGTGCTGAGATCTTTAGTGGCTCCA 1959
DB 106 GACACAGAGGTTGGTTAAATTGACAGTACGTGCTGAGATCTTTAGTGGCTCCA 47
QY 1960 TTGGACAGAGAGCCGGAAGTCCATATCGGTAACAAGTGTGGCC 2004
DB 46 TTGGACAGAGAGCCGGAAGTCCATATCGGTAACAAGTGTGGAC 2

RESULT 5
US-09-385-982-300/c
; Sequence 300, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS: II
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 300
; LENGTH: 591
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(591)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-300

Query Match
Best Local Similarity 8.5%; Score 284.8; DB 4; Length 591;
Matches 286; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2480 TGATTGATTAATTTAGAGCTGTGTATCCGATTAAGAAAGATTAAGCAAGAT 2539
DB 289 TCAATGATTAATTTAGAGCTGTGTATCCGATTAAGAAAGATTAAGCAAGAT 230
QY 2540 ATGTTGAAGTGTCAAGATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAGAA 2599
DB 229 ATGTTGAAGTGTCTCAAGATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAGAA 170
QY 2600 TCTTTGAATTTATATGCAAGTCTATTTCAGCAACACCATCTCATCTATTACTTTTC 2659
DB 169 TGTTTGAATTTATATGCAAGTCTATTTCAGCAACACCATCTCATCTATTACTTTTC 110
QY 2660 ATCTAAGTGCATTAATTTTAAACAGATATTCCTGCTGCTTAAATTTGCTA 2719
DB 109 ATCTAAGTGCATTAATTTTAAACAGATATTCCTGCTGCTTAAATTTGCTA 50
QY 2720 AATATTTCTTTTGTAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2768
DB 49 AATATTTCTTTTGTAGTGTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1

RESULT 6
US-09-301-665-3
; Sequence 3, Application US/09301665
; Patent No. 6207876
; GENERAL INFORMATION:
; APPLICANT: KELLEMS, RODNEY E.
; APPLICANT: DATTA, SURJIT K.
; APPLICANT: BLACKBURN, MICHAEL R.
; TITLE OF INVENTION: ADENOSINE DEAMINASE DEFICIENT TRANSGENIC MICE AND
; TITLE OF INVENTION: METHODS FOR THE USE THEREOF
; FILE REFERENCE: UTS-243
; CURRENT APPLICATION NUMBER: US/09/301,665
; CURRENT FILING DATE: 1999-04-28
; EARLIER APPLICATION NUMBER: 60/083,408
; EARLIER FILING DATE: 1998-04-29
; EARLIER APPLICATION NUMBER: 60/083,370
; EARLIER FILING DATE: 1998-04-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 36741
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-301-665-3

Query Match
Best Local Similarity 5.7%; Score 191.8; DB 4; Length 36741;
Matches 217; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 2717 CTAAATATTTCTTTTGTAGTGTGAGTCTTGTCTGTGCGCCAGGCTGAGTACAGTGT 2776
DB 19604 CTTTCTTTTCTTTTGTAGTGTGAGTCTTGTCTGTGCGCCAGGCTGAGTACAGTGT 19663
QY 2777 GTGATCCAGCTCAGCACTCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2836
DB 19664 ATGATCTCGCTCAGCACTCCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 19723
QY 2837 TCTTAAGTGTGCTTTTACAGGACCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2896
```

```
Db      19724 TCCGAGTACTGAGATTACAGACATGCGCACCATCCAGCTAATTTGATTTT 19783
QY      2897 ATAGAGACGGGTTTCCGATTGGCCAGGCTGTCTTGAATCTTACGCTCAAGTGATC 2956
Db      19784 GCGGAGATGGGGTTTTCATGTTGGCCAGGCTGTGTCTGAACTCTTCACTCAAGTATC 19843
QY      2957 TGCCTGCTGTGTCTTCCCA 2975
Db      19844 CACCCAGCTTGGCTCCCA 19862

RESULT 7
US-09-345-882-1
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bouquelerec, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103806
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A

; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108308
; OTHER INFORMATION: 5-135-357 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108471
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134134
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134362
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134374
; OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146328
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
; FEATURE:
```


D _b	3080	ACAGGCGCACATGCCTGGTGTTTTGTTTTGTTTTTTTTTTTGAAGACTGAGTCCTGCT	3139
Q _y	2750	CTGTGGCCCAAGCGGTGAGTAACAGTGTATATCCAGCTCACTGAACTCCTGGCTCTGT	2809
D _b	3140	CTGTGGCTTAGCGGTGAGTGTGACAGTGCATGATCTGGGTCAACCGAACCTTGTGCTCTGG	3199
Q _y	2810	GATTACATGATATTTCTCTGCTCTCAGCTTCTCTAAGTAGCTGGGATTACAGGACCACCA	2869
D _b	3200	GGTTCAAGCGATTTCTCTGCTTAGCCCTCTTAAGTAGCTGGGATTTACAGGACCGACAC	3259
Q _y	2870	CATGCCCAAGCTAATTTTGTGTAATTTTAAATAAGACGGGGTTTCGCCAATTTGGCCNAGGTG	2929
D _b	3260	CATGCCCGGCTAATTTTGTATTTGTATTAAGAACAAGGGTTTTTCCTGTTGGCCAGGCTG	3319
Q _y	2930	GTTTGAAGTCTCTGACGTCAGTGTATCGCTGGCTGGTCTGCCAA	2976
D _b	3320	GTCCTGAAGTCTCTGACTTCAAGTTATCTGCCCACTCGGCTCTCCAA	3366

```

RESULT 9
US-08-520-373D-4
: Sequence 4, Application US/08520373D
: Patent No. 6451763
: GENERAL INFORMATION:
: APPLICANT: Tombrian-Tink, Joyce
: APPLICANT: Steele, Fintan R
: APPLICANT: Chader, Gerald J
: APPLICANT: Becerra, Sofia P
: APPLICANT: Johnson, Lincoln V
: APPLICANT: Rodriguez, Ignacio R
: TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIUM DERIVED NEUROTROPIC FACTOR
: FILE REFERENCE: 2026-4203US1
: CURRENT APPLICATION NUMBER: US/08/520,373D
: CURRENT FILING DATE: 1995-08-29
: PRIOR APPLICATION NUMBER: 08/377,710
: PRIOR FILING DATE: 1995-01-25
: PRIOR APPLICATION NUMBER: 08/379,979
: PRIOR FILING DATE: 1994-07-25
: PRIOR APPLICATION NUMBER: 07/694,215
: PRIOR FILING DATE: 1992-06-04
: PRIOR APPLICATION NUMBER: 07/952,796
: PRIOR FILING DATE: 1992-09-24
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 14581
: TYPE: DNA
: ORGANISM: HUMAN
: FEATURE:
: OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
: OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
: OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
: US-08-520-373D-4

```

	Query Match	5.7%	Score 190.8	DB 4	Length 14581
	Best Local Similarity	82.7%	Pred. No. 1,le=42		
	Matches 230	Conservative 0	Mismatches 47	Indels 1	Gaps 1
QY	2700	TTGTCCTTAATATTGCTAAATAATTCCTTTTGGAGTGAGTCTCTGTGCGCCA	2759		
Db	9505	TTCTACATTTTCTTTCTTTTCTTTTCTTTTGGAGATGAGTCTGCTGTGCGCCA	9564		
QY	2760	GGCTGGAGTACAGAGGTGTGATATCCAGTCACTGCACTCGGCTCTGGGTTCAATG	2819		
Db	9565	GGCTGGAGTACAGAGGTGTGATATCCAGTCACTGCACTCGGCTCTGGGTTCAATG	9624		
QY	2820	ATTCTCCCTGACCTCTCTAAGTGTGGTTTACAGGACCCACCAACCATGCGCCAGC	2879		
Db	9625	ATTCTCCCTGACCTCTCTAAGTGTGGATTTACAGGACGTGCGCCACCAACCCGCG	9684		
QY	2880	TAAATTTTGTATTTTAAATAGAGACGGGGTTTCGCAT-TTGGCCAGAGCTGTCTTGAAC	2938		
Db	9685	TAAATTTTGTATTTTAAATAGAGAGGGGGTTTCTCAGTTGGCCAGGATGATCTCAAC	9744		

Qy 2339 TCTGAGCTCAAGTGAATCGCTGCGCTTGTCTCCAA 2376
| | | | | | | | | | | | | | | | | | | | | |
Db 9745 TCCTGACCTCAAGTGAATCTACCCGCGCTCGGCTCTCAA 9782

```

RESULT 10
US-08-367-841A-43
Sequence 43, Application US/08367841A
Patent No. 6319687
GENERAL INFORMATION:
APPLICANT: Chader, Gerald J.; Rodriguez,
APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
APPLICANT: Tombran-Tink, Joyce
TITLE OF INVENTION: PIGMENT BETHLELUM
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PDEF GENE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Morgan & Finnegan
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatibld
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/367,841A
FILING DATE: 30-DEC-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,963
FILING DATE: 07-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 22481 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: Pl-147
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: full length genomic
OTHER INFORMATION: sequence for PDEF plus flanking sequences
US-08-367-841A-43

```

	Query Match	5.7%	Score 190.8	DB 4	Length 22481
	Best Local Similarity	82.7%	Pred. No. 1.4e-42		
	Matches 230	Conservative 0	Mismatches 47	Indels 1	Gaps 1
QY	2700	TTTGCCTTAAATTGCTAAATATCTTTTGGAGTGGAGCTTCTCTGCGCCA			2759
DB	9494	TTTACATTTTTTTTCCTTTTTTTTTTTTTTTTGGATGGAGCTCTGTGCGCCA			9553
QY	2760	GGCTGAGTACAGTGTGTGATCCACGTCACCACTCCGCTCTGGGATTCACATG			2819
DB	9554	GCTTGAGTGTGAGTGGCGCATCTCAGCTACCTGCACACTTCTCTGGATTCAGTG			9613

Accession	Sequence	Position
QY	2820 TTCTCCGCGCTCAGGCTCTCTAAGTAGGTGGGGTTACAGGGACCAACACATGCGCAG	2879
Db	9614 ATTCCTCGCTTAGGCTCCCAAGTAGGTGGGATTAACGGATGGGCACCAACCGGGC	9673
QY	2880 TAATTTTGTATTTTAAATAGAGACGGGGTTTGGCCAT -TTGGCAGGCTGCTTTGAC	2938
Db	9674 TAAATTTTGTATTTTAAATAGAGATGGGGTTTCTCCAGTTGGCCAGGATGCTCAAC	9733
QY	2939 TCTGACGTCAGTGATCTGCCTGCGCTTGATCCCA	2976
Db	9734 TCTGACCTCAGGTGATTAACCGGCTCGGCTCTCAA	9771

```

RESULT 11
PCT-US95-07201-43
Sequence 43: Application PC/TUS9507201
GENERAL INFORMATION:
APPLICANT: Chader, Gerald J.; Becerra, Sofia
APPLICANT: Patricia; Schwartz, Joan P.;
APPLICANT: Tanikawa, Takayuki
TITLE OF INVENTION: PIGMENT EPITHELIUM
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Flinnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07201
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/367,841
FILING DATE: 30-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/257,863
FILING DATE: 07-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/952,796
FILING DATE: 24-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: DOROTHY R. AUTH
REGISTRATION NUMBER: 36434
REFERENCE/DOCKET NUMBER: 20264126PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 22481 Base Pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Unknown
MOLECULE TYPE: Genomic DNA
FEATURE:
NAME/KEY: P1-147
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION: full length genomic
OTHER INFORMATION: sequence for PEDF plus flanking sequences.
PCT-US95-07201-43

```

	Best Local Similarity	82.7%	Pred No. 1,4e-42;	Matches 230;	Conservative 0;	Mismatches 47;	Indels 1;	Gaps 1;
QY	2700	TTGTCCTTAATATTGCTAAATATTTCTTTTGTAGGTGAGCTTGTCTCTGTGGCCCA	2759					
Db	9494	TTCTACATTTTCTTTTCTCTTTTCTTTTCTTTTGTGATGGAGCTGCTCTGTGGCCCA	9553					
QY	2760	GGCGGAGTACAGTGGGTATCCCACTACATCGCAACCTCCGCTCTGGTTCACATG	2819					
Db	9554	GCTGGAGTACAGTGGGCGCATCTCACTACATCGCAACCTCCATTTCTGGATTCAAGTG	9613					
QY	2820	ATTCTCCGTCCAGCTCTCTAAAGTAGCTGGGTATTACAGGCACCCACCACTGCGCAGC	2879					
Db	9614	ATTCTCCGCTTAGCTCCCAAGTAGCTGGGATTAAGGCAATGCGCACCAACACCGGCG	9673					
QY	2880	TAAATTTTGTATTTTAAATAGACAAGGGGTTTGCCAT-TTGGCAGGCTGTCTTGAAC	2938					
Db	9674	TAAATTTTGTATTTTGTAGAGATGGGGTTTCCCATGTGGGCAAGATGTCTCAAAC	9733					
QY	2939	TCCGAGCTCAGATGTGCTGGCCCTTGGATGCCAA	2976					
Db	9734	TCTGACCTCAGGTGATCTAACCGGCTTCGGCTCTCAA	9771					

```

RESULT 12
US-09-875-223-2
Sequence 2, Application US/09875223
Patent No. 6391850
GENERAL INFORMATION:
APPLICANT: No. 6391850 Northwestern University
APPLICANT: No. 63918501 Bouck
APPLICANT: David Dawson
APPLICANT: Paul Gillis
TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis
FILE REFERENCE: 0290-2303
CURRENT APPLICATION NUMBER: US/09/875,223
CURRENT FILING DATE: 2001-06-06
PRIORITY APPLICATION NUMBER: US 09/122,079
PRIORITY FILING DATE: 1998-07-23
PRIORITY APPLICATION NUMBER: PCT/US98/15228
PRIORITY FILING DATE: 1998-07-23
PRIORITY APPLICATION NUMBER: US 08/899,304
PRIORITY FILING DATE: 1997-07-23
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 22484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Unsure
LOCATION: 1..22484
OTHER INFORMATION: "n" means either a, c, t, or g
US-09-875-223-2

```

	Query Match	5.7%;	Score 190.8;	DB 42;	Length 22484;
	Best Local Similarity	82.7%;	Prod. No. 1.4e-22;		
	Matches 230;	Conservative 0;	Mismatches 47;	Indels 1;	Gaps 1;
QY	2700	TTTGCTTTAAATATTGCTAAATATTTCTTTTGAAGTGGAGCTCTGCTGTGCGCCA	2759		
Db	9494	TTCTACATTTTTTTTTTCCCTTTTTTTTTTTTTTGGATGAGATCTGTCTGTGGCCA	9553		
QY	2760	GGCTGAGTACAGTGTGTGATCCACGACTCACTGCACACTCCGCTCTGGGTTCACATG	2819		
Db	9554	GCCTGGAGTCAAGTGGGGCGCATCTCAGCTACGTCAACACGCCACTTCTGGATTCAAGTG	9613		
QY	2820	ATTCTCTCGCTCAGCTTCTCTAAGTACCTGGGTTTACAGGCACCCACACATGAGCCAGC	2879		
Db	9614	ATTCTCTCGCTTACGCTCCCAAGTACGTGGGATTCACAGGCATGCGCACACACCCGGC	9673		
QY	2880	TAAATTTTGTATTTTAAATATGAGACGGGGTTTGCCAT-TTGGCCACAGCTGTCTTGAAC	2938		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 21, 2003, 21:12:29 / Search time 447 Seconds
(without alignments)
10981.083 Million cell updates/sec

Title: US-10-025-380-1076

Perfect score: 3345
Sequence: 1 gaattccgctgcacacatg.....cctttgcgcgcgaattc 3345

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PTCUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3345	100.0	3345	9	US-10-025-380-1076
2	3345	100.0	3345	10	US-09-922-217-1076
3	3345	100.0	3345	10	US-09-862-436-302
4	3345	100.0	3345	10	US-09-833-263-1076
5	3345	100.0	3345	10	US-09-880-107-3288
6	3314.8	99.1	3654	10	US-09-962-832-118
7	3314.8	99.1	3654	10	US-09-880-107-3807
8	2432.6	72.7	2877	9	US-10-025-380-1086
9	2432.6	72.7	2877	10	US-09-922-217-1086
10	2432.6	72.7	2877	10	US-09-833-263-1086
11	514.8	15.4	518	10	US-09-998-598-1266
12	514.2	15.4	520	9	US-10-046-935-1019
13	514.2	15.4	520	9	US-09-878-178-1019
14	514.2	15.4	520	9	US-10-146-502-1019
15	512.4	15.3	517	9	US-10-066-543-3247
16	511.4	15.3	627	9	US-09-871-161-514
17	511	15.3	512	9	US-10-066-543-2636
18	510.4	15.3	512	9	US-10-066-543-2512
19	508.8	15.2	512	9	US-10-060-036-2188

20	497	14.9	498	9	US-10-046-935-2092	Sequence 2092, Ap
21	497	14.9	498	9	US-09-878-178-2092	Sequence 2092, Ap
22	497	14.9	498	9	US-10-146-502-2092	Sequence 2092, Ap
23	490.4	14.4	496	9	US-10-060-036-2093	Sequence 2093, Ap
24	479.8	14.3	495	9	US-10-060-036-1950	Sequence 1950, Ap
25	469.2	14.0	492	9	US-10-046-935-571	Sequence 571, App
26	469.2	14.0	492	9	US-09-878-178-571	Sequence 571, App
27	469.2	14.0	492	9	US-10-146-502-571	Sequence 571, App
28	461.4	13.8	489	9	US-10-046-935-98	Sequence 98, App1
29	461.4	13.8	489	9	US-09-878-178-98	Sequence 98, App1
30	461.4	13.8	489	9	US-10-146-502-98	Sequence 98, App1
31	450	13.5	450	10	US-09-998-598-696	Sequence 696, App
32	450	13.5	450	10	US-09-998-598-1002	Sequence 1002, App
33	445.2	13.3	450	10	US-09-998-598-1285	Sequence 1285, App
34	431.4	12.9	516	10	US-09-920-300A-409	Sequence 409, App
35	431.4	12.9	516	12	US-10-033-528-409	Sequence 409, App
36	425.6	12.7	451	9	US-10-066-543-1412	Sequence 1412, Ap
37	420.2	12.6	451	9	US-10-066-543-1413	Sequence 1413, Ap
38	410.4	12.3	568	10	US-09-920-300A-705	Sequence 705, App
39	410.4	12.3	568	12	US-10-033-528-705	Sequence 705, App
40	408.4	12.2	568	10	US-09-998-598-843	Sequence 843, App
41	401.4	12.0	429	9	US-10-066-543-1531	Sequence 1531, App
42	386.4	11.6	568	10	US-09-998-598-723	Sequence 723, App
43	367.4	11.0	375	9	US-10-066-543-2626	Sequence 2626, App
44	367	11.0	367	9	US-10-060-036-1960	Sequence 1960, Ap
45	367	11.0	367	9	US-10-060-036-2398	Sequence 2398, Ap

ALIGNMENTS

RESULT 1
US-10-025-380-1076
Sequence 1076, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yudu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1076
LENGTH: 3345
TYPE: DNA
ORGANISM: Homo sapiens
US-10-025-380-1076

Query Match 100.0%; Score 3345; DB 9; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GAATTCGCTGCACCACTGAATGAAAGAAAGCTTTTAACACCACTTTGTGACTTA 60
DB 1 GAATTCGCTGCACCACTGAATGAAAGAAAGCTTTTAACACCACTTTGTGACTTA 60

QY 61 CAGAAAGAAATTTGATTAAGAAAACTATGATCTTCAGAGCCCATCTTCACTCCCTGT 120
| | | | |
Db 61 CAGAAAGAAATTTGATTAAGAAAACTATGATCTTCAGAGCCCATCTTCACTCCCTGT 120
QY 121 CTTCTTAATGCTTTATTTGGCACTGGAATATGGCCAAAGGGGAAGTTTATGAGCCCTG 180
| | | | |
Db 121 CTTCTTAATGCTTTATTTGGCACTGGAATATGGCCAAAGGGGAAGTTTATGAGCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATGAAAGCCCAAGACCCGAGTCAAAATTAATTCAGATT 240
| | | | |
Db 181 AAACCCATGACATTTTCTATTTATGAAAGCCCAAGACCCGAGTCAAAATTAATTCAGATT 240
QY 241 AAGCCCAATCTCTGCTGTGATCTTTGAATTAATGAGGAGACAGACAAATATTTGTG 300
| | | | |
Db 241 AAGCCCAATCTCTGCTGTGATCTTTGAATTAATGAGGAGACAGACAAATATTTGTG 300
QY 301 ATAGAACGGAGGAGACTTGTATTAACAACAGACCTTTGACAGAGGAAACAAGATCTACT 360
| | | | |
Db 301 ATAGAACGGAGGAGACTTGTATTAACAACAGACCTTTGACAGAGGAAACAAGATCTACT 360
QY 361 CACATCTCCAGGTTGACAGCCCTGAGCTTAATGGAATTAATGAGAGGCTCCAGTCCCT 420
| | | | |
Db 361 CACATCTCCAGGTTGACAGCCCTGAGCTTAATGGAATTAATGAGAGGCTCCAGTCCCT 420
QY 421 ATCACCATAGAAAGTGAAGGACATCAACGACATGACCCAGTTTCTCCAGTCAAAAGTAC 480
| | | | |
Db 421 ATCACCATAGAAAGTGAAGGACATCAACGACATGACCCAGTTTCTCCAGTCAAAAGTAC 480
QY 481 GAAAGCTCAGTAAGGAGGACATCTGCGCCAGAGAAAGCCCTTCTTGTATGTCAATGACACA 540
| | | | |
Db 481 GAAAGCTCAGTAAGGAGGACATCTGCGCCAGAGAAAGCCCTTCTTGTATGTCAATGACACA 540
QY 541 GACCTGATGATTCGGGCACTCCCAATGGCCAGCTTTATTAACGATTTGTATCCACTT 600
| | | | |
Db 541 GACCTGATGATTCGGGCACTCCCAATGGCCAGCTTTATTAACGATTTGTATCCACTT 600
QY 541 GACCTGATGATTCGGGCACTCCCAATGGCCAGCTTTATTAACGATTTGTATCCACTT 600
| | | | |
Db 541 GACCTGATGATTCGGGCACTCCCAATGGCCAGCTTTATTAACGATTTGTATCCACTT 600
QY 601 CCCATGATCAACAATGTCATGATCTTTCAGATCAACAACAAGGGAGCCATCTCTT 660
| | | | |
Db 601 CCCATGATCAACAATGTCATGATCTTTCAGATCAACAACAAGGGAGCCATCTCTT 660
QY 661 ACCGAGAGGGGATCTCAGGAATTAATCTGCTAAGAAATCTTCTAATATCTGATATC 720
| | | | |
Db 661 ACCGAGAGGGGATCTCAGGAATTAATCTGCTAAGAAATCTTCTAATATCTGATATC 720
QY 721 TCAATGAAAGGACATGGGAGGCCAGATGGAATTCCTTCAGTGAATCAACAATCTGTGAT 780
| | | | |
Db 721 TCAATGAAAGGACATGGGAGGCCAGATGGAATTCCTTCAGTGAATCAACAATCTGTGAT 780
QY 781 ATCATAGTGAAGAATAATTTGGAAAGCACCAAAACCTGTGAGATGTGGAATACTCA 840
| | | | |
Db 781 ATCATAGTGAAGAATAATTTGGAAAGCACCAAAACCTGTGAGATGTGGAATACTCA 840
QY 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCGGGTGCAAAATAT 900
| | | | |
Db 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGCGGTGAATGATCCGGGTGCAAAATAT 900
QY 901 TCCTTAATGTTGCAAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTTGACAGAGAGAT 960
| | | | |
Db 901 TCCTTAATGTTGCAAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTTGACAGAGAGAT 960
QY 961 ATTTACGTCATCAGCCCTTGAGACCGAGAGAAAAAGATGATATGTTTTTATGACATT 1020
| | | | |
Db 961 ATTTACGTCATCAGCCCTTGAGACCGAGAGAAAAAGATGATATGTTTTTATGACATT 1020
QY 1021 GCAAAAGATGAGTACGGAATAACAATTTCAATTCGCTGGAATTCATGTAATAAGTTAAA 1080
| | | | |
Db 1021 GCAAAAGATGAGTACGGAATAACAATTTCAATTCGCTGGAATTCATGTAATAAGTTAAA 1080
QY 1081 GATATTAATGATTAATCCACTCAATGTCTGTCACAGTAACCGTAATTTGAGAGTCCAGAG 1140
| | | | |
Db 1081 GATATTAATGATTAATCCACTCAATGTCTGTCACAGTAACCGTAATTTGAGAGTCCAGAG 1140
QY 1141 AATGAAGACCTGGGTAAACGTAATGGGAGCCCTTACTGACATGACAGGAGATGAAGAAAT 1200
| | | | |

Db 1141 AATGAAGACCTGGGTAAACGTAATGGGAGCCCTTACTGACATGACAGGATGAAGAAAT 1200
| | | | |
QY 1201 ACTGCCAACAGTTTCTTAAACTACAGATTTGGAGACAACTCCCAAACTTCCCATGAT 1260
| | | | |
Db 1201 ACTGCCAACAGTTTCTTAAACTACAGATTTGGAGACAACTCCCAAACTTCCCATGAT 1260
QY 1261 GGACTCTTCTTAATCCAAACCTATGCTGGAATGTATCAGTTAGCTAAACAGTCTTGAAG 1320
| | | | |
Db 1261 GGACTCTTCTTAATCCAAACCTATGCTGGAATGTATCAGTTAGCTAAACAGTCTTGAAG 1320
QY 1321 AAGCAAGATACCTCTCAGTACAACTTAACGATAGAGTGTCTGACAAAGATTTCAAGACC 1380
| | | | |
Db 1321 AAGCAAGATACCTCTCAGTACAACTTAACGATAGAGTGTCTGACAAAGATTTCAAGACC 1380
QY 1381 CTTTGTGTTGTGCAAAATCAACGTTATGATATCAATGATCAAGTCCCATCTTGAAGAAA 1440
| | | | |
Db 1381 CTTTGTGTTGTGCAAAATCAACGTTATGATATCAATGATCAAGTCCCATCTTGAAGAAA 1440
QY 1441 TCAGATTAATGGAACCTGACTCTTGTGGAAGACAAACATTTGGTCCACATCTTAACC 1500
| | | | |
Db 1441 TCAGATTAATGGAACCTGACTCTTGTGGAAGACAAACATTTGGTCCACATCTTAACC 1500
QY 1501 ATCCAGGCCACTGATCTGATGAGCCATTAATCTGGAGTTCTTAATAATCTGTATCATATC 1560
| | | | |
Db 1501 ATCCAGGCCACTGATCTGATGAGCCATTAATCTGGAGTTCTTAATAATCTGTATCATATC 1560
QY 1561 ATTAAGGAGACAGTGAAGGAGCCCTGGGGTGTGAACAAGATCCCATACCAACCCGA 1620
| | | | |
Db 1561 ATTAAGGAGACAGTGAAGGAGCCCTGGGGTGTGAACAAGATCCCATACCAACCCGA 1620
QY 1621 TATGTCAATATTAAGAACCTCTTGAATTTGAAACAGACGCTTTTCAACATTTGTGTC 1680
| | | | |
Db 1621 TATGTCAATATTAAGAACCTCTTGAATTTGAAACAGACGCTTTTCAACATTTGTGTC 1680
QY 1681 AAGCAGAAAAATCTGAGCCCTAGTGTGTTGTGTGAAGTACAAATCAAGTTCTTTGCC 1740
| | | | |
Db 1681 AAGCAGAAAAATCTGAGCCCTAGTGTGTTGTGTGAAGTACAAATCAAGTTCTTTGCC 1740
QY 1741 AAGTTCAAGCTTATTTGTGACAGATGGAATGAAGACCTTCAATTTTCCCAACCGTATTC 1800
| | | | |
Db 1741 AAGTTCAAGCTTATTTGTGACAGATGGAATGAAGACCTTCAATTTTCCCAACCGTATTC 1800
QY 1801 CAAGCGAAAGTCAAGTGAAGATGATGATGAGCACTAAGATGGGCAATGATCTGCCAAG 1860
| | | | |
Db 1801 CAAGCGAAAGTCAAGTGAAGATGATGATGAGCACTAAGATGGGCAATGATCTGCCAAG 1860
QY 1861 GATCAGAAAGTCTGACATTAAGCTATTCATCTGAGGGGAGACAAAGAGTTGGCTTAAA 1920
| | | | |
Db 1861 GATCAGAAAGTCTGACATTAAGCTATTCATCTGAGGGGAGACAAAGAGTTGGCTTAAA 1920
QY 1921 ATTGACCAAGTGAAGTGTGATCTTTAATGTGTGCTTCAATTTGACAGAGAGCGGAAGT 1980
| | | | |
Db 1921 ATTGACCAAGTGAAGTGTGATCTTTAATGTGTGCTTCAATTTGACAGAGAGCGGAAGT 1980
QY 1981 AATGACCAAGTGAAGTGTGATCTTTAATGTGTGCTTCAATTTGACAGAGAGCGGAAGT 1980
| | | | |
Db 1981 AATGACCAAGTGAAGTGTGATCTTTAATGTGTGCTTCAATTTGACAGAGAGCGGAAGT 1980
QY 1981 CCATATCGGGTAAAGAGTGTGAGCAAGAAAGTGAAGGAGGAGGCTTCTTAAGCTCTGTGTC 2040
| | | | |
Db 1981 CCATATCGGGTAAAGAGTGTGAGCAAGAAAGTGAAGGAGGAGGCTTCTTAAGCTCTGTGTC 2040
QY 2041 GAGTTCCACTGATCTTATGATGATGATGATCAACCTTCCAGGCTAAGCAAGACTAC 2100
| | | | |
Db 2041 GAGTTCCACTGATCTTATGATGATGATGATGATCAACCTTCCAGGCTAAGCAAGACTAC 2100
QY 2101 ACGGGCTTGTCTTCTGCACTCCCTCAAGTGCACCTGGAAGTCTATTTTCCAGGCTACT 2160
| | | | |
Db 2101 ACGGGCTTGTCTTCTGCACTCCCTCAAGTGCACCTGGAAGTCTATTTTCCAGGCTACT 2160
QY 2161 GATGATGATGAGCACTTAATTTGGGGGCTCCCATTTTAATTTTCCCTGGCAGTGAAGC 2220
| | | | |
Db 2161 GATGATGATGAGCACTTAATTTGGGGGCTCCCATTTTAATTTTCCCTGGCAGTGAAGC 2220
QY 2221 TTACAAAACGACTGGGAAGTTTCAAAATCAATGATGATGATGATGATGATGATGATGATGAT 2280
| | | | |

Db 2221 TTACAAAACGACTGGAACTTTCCAAAATCAATGTACTATGCCCACTGTCTACAGG 2280
Qy 2281 CACACAGACTTGGAGAGAGGCGTATGCTCTTGAATCCGATCAATGATGGGGTCCG 2340
Db 2281 CACACAGACTTGGAGAGAGGCGTATGCTCTTGAATCCGATCAATGATGGGGTCCG 2340
Qy 2341 CCACCCCTTGGAGAGGCAATGTTCTTTTACCACTTCAATTCGAGTGTGTGGAAGAGT 2400
Db 2341 CCACCCCTTGGAGAGGCAATGTTCTTTTACCACTTCAATTCGAGTGTGTGGAAGAGT 2400
Qy 2401 TGTTCGGGCGAGAGGTCACACAGCTGGGATACCCCTGGGCGATGGGAGTGTGTATA 2460
Db 2401 TGTTCGGGCGAGAGGTCACACAGCTGGGATACCCCTGGGCGAGTGGGAGTGTGTATA 2460
Qy 2461 CTGCTGACCAACCTTCTGTGATGATGATTAATTTTACAGATGTGTATCCGATTAAG 2520
Db 2461 CTGCTGACCAACCTTCTGTGATGATGATTAATTTTACAGATGTGTATCCGATTAAG 2520
Qy 2521 AAGGATTAAGGCAAGATATGTTGAAAGTCTCAGCATCTGAAGTCAAACTCTGAGA 2580
Db 2521 AAGGATTAAGGCAAGATATGTTGAAAGTCTCAGCATCTGAAGTCAAACTCTGAGA 2580
Qy 2581 AGCGTAATTTGAAAGAAATGTTGATTTATATAGCAAGTGTATTCAGCAACCA 2640
Db 2581 AGCGTAATTTGAAAGAAATGTTGATTTATATAGCAAGTGTATTCAGCAACCA 2640
Qy 2641 TCTCATCTTATCTTCTTCTTCACTTCAAGTCAATTAATTTTAAAGATATTCCTCT 2700
Db 2641 TCTCATCTTATCTTCTTCTTCACTTCAAGTCAATTAATTTTAAAGATATTCCTCT 2700
Qy 2701 TGTCTTTTAAATTTTGTAAATATTTCTTTTGGAGTGAAGTCTGCTGTGCCAG 2760
Db 2701 TGTCTTTTAAATTTTGTAAATATTTCTTTTGGAGTGAAGTCTGCTGTGCCAG 2760
Qy 2761 GCTGGAATAGAGGAGTGTATCCAGCTCACTGACACTCCGCTCTGTGGTTCATGA 2820
Db 2761 GCTGGAATAGAGGAGTGTATCCAGCTCACTGACACTCCGCTCTGTGGTTCATGA 2820
Qy 2821 TTTCTCTGCTCAGCTTCTTAAGTGTGTGTATAGAGGACCCACCACTATGCCAGCT 2880
Db 2821 TTTCTCTGCTCAGCTTCTTAAGTGTGTGTATAGAGGACCCACCACTATGCCAGCT 2880
Qy 2881 AATTTTGTATTTTAAATAGAGCGGGATTTGCCATTTGGCCAGGCTGTGTGAATC 2940
Db 2881 AATTTTGTATTTTAAATAGAGCGGGATTTGCCATTTGGCCAGGCTGTGTGAATC 2940
Qy 2941 CTGACGTCAGATGATCTGCTGCTGTGTGTCTCCCAATACAGGCAATGACACCA 3000
Db 2941 CTGACGTCAGATGATCTGCTGCTGTGTGTCTCCCAATACAGGCAATGACACCA 3000
Qy 3001 CCTACTAGATATTTCTGATGTATAGACATTAAGAGATTTTTCATTTTCCATGACAT 3060
Db 3001 CCTACTAGATATTTCTGATGTATAGACATTAAGAGATTTTTCATTTTCCATGACAT 3060
Qy 3061 TTTTCTCTCTGCAAAATGAGCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Db 3061 TTTTCTCTCTGCAAAATGAGCTTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Qy 3121 CATTAATATTTCTGATATTTTCTTTATCAAGAGATATATCAAGTGTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATATTTTCTTTATCAAGAGATATATCAAGTGTGTCTCATAGA 3180
Qy 3181 ACTGCTGATTCATTTATGTTTCTTATCAATCTGATGCTGCTGCTGCTGCTGCTGCTG 3240
Db 3181 ACTGCTGATTCATTTATGTTTCTTATCAATCTGATGCTGCTGCTGCTGCTGCTGCTG 3240
Qy 3241 TCTTTGTATTTCACTGAATTTCAACATTTGTCAAGAGAGAGAGAGAGAGAGAGAG 3300
Db 3241 TCTTTGTATTTCACTGAATTTCAACATTTGTCAAGAGAGAGAGAGAGAGAGAGAG 3300
Qy 3301 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3345
Db 3301 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3345

RESULT 2
US-09-922-217-1076
Sequence 1076, Application US/09922217
Patent No. US2002007641A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tonglong
APPLICANT: Jiang, Yuguang
APPLICANT: Smith, Carole Lynn
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C13
CURRENT APPLICATION NUMBER: US/09/922,217
NUMBER OF FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 1124
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1076
LENGTH: 3345
TYPE: DNA
ORGANISM: Homo sapiens
US-09-922-217-1076

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAATTCCTGCTCGACCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1 GAATTCCTGCTCGACCATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 61 CAGAAAGAAATTTGAAATTAAGAAACTATGATATCTTACGGCCATCTTCACTCCGTGT 120
Db 61 CAGAAAGAAATTTGAAATTAAGAAACTATGATATCTTACGGCCATCTTCACTCCGTGT 120
Qy 121 CTCTTATGCTTATTTGGCAATGATATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 CTCTTATGCTTATTTGGCAATGATATGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 181 AAACCATGACATTTTCTATTTATGAAAGCGCAAGACGAGTCAATTTATTTCCAGTTT 240
Db 181 AAACCATGACATTTTCTATTTATGAAAGCGCAAGACGAGTCAATTTATTTCCAGTTT 240
Qy 241 AAGGCCAATCTCTCTGCTGATTTTGAATTAAGCTTGGAGAGACAGAAATATTTACT 300
Db 241 AAGGCCAATCTCTCTGCTGATTTTGAATTAAGCTTGGAGAGACAGAAATATTTACT 300
Qy 301 ATGAAAGGAG 360
Db 301 ATGAAAGGAG 360
Qy 361 CACAACTCTCAGGTTGACAGCCCTGAGAGCTTAATGAAATTAATGAGAGAGAGAGAG 420
Db 361 CACAACTCTCAGGTTGACAGCCCTGAGAGCTTAATGAAATTAATGAGAGAGAGAGAG 420
Qy 421 ATCACCATGAG 480
Db 421 ATCACCATGAG 480
Qy 481 GAAAGCTCAGTAAAG 540
Db 481 GAAAGCTCAGTAAAG 540
Qy 541 GACCTGATGATCCGCGCACTCCCAATGAGCCAGCTTATTAACAGATGTCTCAGCTT 600

Db 541 GACCTGGATGATCGGCGCACTCCCAATGGCGAGCTTTATTAACGATTTGTCATCAGCTT 600
Qy 601 CCCATGATCAACAATGTCATGTAATCTTTCAGATCAACAACAAACGGAGCCATCTCTT 660
Db 601 CCCATGATCAACAATGTCATGTAATCTTTCAGATCAACAACAAACGGAGCCATCTCTT 660
Qy 661 ACCCGAAGGGATCTCAGGAATGTAATCCCTGTAAGATCCCTCTTAATCTGGTATC 720
Db 661 ACCCGAAGGGATCTCAGGAATGTAATCCCTGTAAGATCCCTCTTAATCTGGTATC 720
Qy 721 TCAAGTGAAGGACATGGAGGCGCAGAGTGAAGTCTTTCAGTGAATACCACTGTGGAT 780
Db 721 TCAAGTGAAGGACATGGAGGCGCAGAGTGAAGTCTTTCAGTGAATACCACTGTGGAT 780
Qy 781 ATCATAGTACAGAGAAATTTTGGAAAGCACAAACCTGTGGAGATGTGGAAACCTCA 840
Db 781 ATCATAGTACAGAGAAATTTTGGAAAGCACAAACCTGTGGAGATGTGGAAACCTCA 840
Qy 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGGGTGGATGATCCCGGTGCAATAT 900
Db 841 ACTGATCTCAACCCCATCAAAATCACTCAGGTGGGTGGATGATCCCGGTGCAATAT 900
Qy 901 TCCTTAGTGAACAAGAGAGAGCTGCCAAGATTTCCATTTTCAATTTGACAGAGAGAT 960
Db 901 TCCTTAGTGAACAAGAGAGAGCTGCCAAGATTTCCATTTTCAATTTGACAGAGAGAT 960
Qy 961 ATTTAGCTGACTCAGCCCTTGGACCGAGAGAGAAAGATGCAATGTTTTTATGCACTT 1020
Db 961 ATTTAGCTGACTCAGCCCTTGGACCGAGAGAGAAAGATGCAATGTTTTTATGCACTT 1020
Qy 1021 GCAAGAGATGATGCGGAGAAACCACTTTCATTCGCGTGGAAATTCATGTAAAGTTAA 1080
Db 1021 GCAAGAGATGATGCGGAGAAACCACTTTCATTCGCGTGGAAATTCATGTAAAGTTAA 1080
Qy 1081 GATATTAATGATTAATCCACTACATATGTCGACACAGTAACCGATTTTGAAGTCCAGAG 1140
Db 1081 GATATTAATGATTAATCCACTACATATGTCGACACAGTAACCGATTTTGAAGTCCAGAG 1140
Qy 1141 AATGAAGAATCGGGTAAACGATATGGGACCTTACTGCAATGACAGGAGATGAAGAAAT 1200
Db 1141 AATGAAGAATCGGGTAAACGATATGGGACCTTACTGCAATGACAGGAGATGAAGAAAT 1200
Qy 1201 ACTGCAACAGTTTCTTAACTACAGATGTTGGAGCAAACTCCCAATCTTCCATGGAT 1260
Db 1201 ACTGCAACAGTTTCTTAACTACAGATGTTGGAGCAAACTCCCAATCTTCCATGGAT 1260
Qy 1261 GGAATCTTCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTTAAACAGTCTTGAAG 1320
Db 1261 GGAATCTTCTTAATCCAAACCTATGCTGGAATGTTACAGTTAGCTTAAACAGTCTTGAAG 1320
Qy 1321 AAGCAAGATCTCTCAGTACACTTAAACGATAGAGGTGTGACAAAGATTTCAAGACC 1380
Db 1321 AAGCAAGATCTCTCAGTACACTTAAACGATAGAGGTGTGACAAAGATTTCAAGACC 1380
Qy 1381 CTGTTGTTGGCAATCAAGCTTATGATTCATGATGATCCCATCTTTGAAAA 1440
Db 1381 CTGTTGTTGGCAATCAAGCTTATGATTCATGATGATCCCATCTTTGAAAA 1440
Qy 1441 TCAGATTAATGGAACCTGACTCTTGGTGAAGACAAACATTTGGGTCCACCATCTTAAAC 1500
Db 1441 TCAGATTAATGGAACCTGACTCTTGGTGAAGACAAACATTTGGGTCCACCATCTTAAAC 1500
Qy 1501 ATCCAGGCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATATC 1560
Db 1501 ATCCAGGCACTGATGCTGATGAGCCATTTACTGGAGTTCTAAATTTCTGTATCATATC 1560
Qy 1561 ATAAAGGAGACATGAGGAGGCGCTGGGGTGAACAGATCCCATACCAACCGGA 1620
Db 1561 ATAAAGGAGACATGAGGAGGCGCTGGGGTGAACAGATCCCATACCAACCGGA 1620
Qy 1621 TATGTCATTAATTAAGGCTCTTGTATTTGAAACAGAGCTGTTTCCAACTGTGTTTC 1680
Db 1621 TATGTCATTAATTAAGGCTCTTGTATTTGAAACAGAGCTGTTTCCAACTGTGTTTC 1680
Db 1621 TATGTCATTAATTAAGGCTCTTGTATTTGAAACAGAGCTGTTTCCAACTGTGTTTC 1680
Qy 1681 AAGCAGAAATCTGAGCCTAGTGTGTTGGTGAAGTACAAATGCAAGTTCTTTGGC 1740
Db 1681 AAGCAGAAATCTGAGCCTAGTGTGTTGGTGAAGTACAAATGCAAGTTCTTTGGC 1740
Qy 1741 AAGTTCAGCTTAATGTCAGATGTAATGAAGACCTCAATTTTCCAAACGATATC 1800
Db 1741 AAGTTCAGCTTAATGTCAGATGTAATGAAGACCTCAATTTTCCAAACGATATC 1800
Qy 1801 CAGCCAAAGTCAGTGAAGTGTAGCTATTTAGGCACTTAAAGTGGCAATGTATGCGCAAG 1860
Db 1801 CAGCCAAAGTCAGTGAAGTGTAGCTATTTAGGCACTTAAAGTGGCAATGTATGCGCAAG 1860
Qy 1861 GATCCAGAAAGTGTGACATTAAGCTATTACCTGAGGGGAGACACAAGAGGTGGCTTAA 1920
Db 1861 GATCCAGAAAGTGTGACATTAAGCTATTACCTGAGGGGAGACACAAGAGGTGGCTTAA 1920
Qy 1921 ATGACCAAGTCAGTGTGAGATCTTTAGTGTGCTCCATTTGACAGAGAGCCGGAAGT 1980
Db 1921 ATGACCAAGTCAGTGTGAGATCTTTAGTGTGCTCCATTTGACAGAGAGCCGGAAGT 1980
Qy 1981 CCATATCGGGTACAGTGTGTGCGCACAGAAATGAGGGGGCTTCTTAAAGCTGTGTCA 2040
Db 1981 CCATATCGGGTACAGTGTGTGCGCACAGAAATGAGGGGGCTTCTTAAAGCTGTGTCA 2040
Qy 2041 GAGTTCACCTGATCTCTTAATGATGTAATGACAAACCTCCAGGCTAGCCAGAGACTAC 2100
Db 2041 GAGTTCACCTGATCTCTTAATGATGTAATGACAAACCTCCAGGCTAGCCAGAGACTAC 2100
Qy 2101 ACGGGCTTGTCTTCTGCGCATCCCTCAGTGCACCTGGAAGTCTATTTGAGGCTACT 2160
Db 2101 ACGGGCTTGTCTTCTGCGCATCCCTCAGTGCACCTGGAAGTCTATTTGAGGCTACT 2160
Qy 2161 GATGATGATAGCATTAATTTGGGGTCCCATTTTACATTTTCCCTCGGCACTGGAAGC 2220
Db 2161 GATGATGATAGCATTAATTTGGGGTCCCATTTTACATTTTCCCTCGGCACTGGAAGC 2220
Qy 2221 TTAACAAAGCACTGGGAGGTTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTAACAAAGCACTGGGAGGTTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
Qy 2281 CACACAGACTTTAGGAGAGAGGCGTATGTCGTTGATCCGATCATGATGATGATGATGAT 2340
Db 2281 CACACAGACTTTAGGAGAGAGGCGTATGTCGTTGATCCGATCATGATGATGATGATGAT 2340
Qy 2341 CCAACCTTGGAGGCAATGTTTCTTAAACAGTTACATTTGCAAGTGTGTGGAAGGAAT 2400
Db 2341 CCAACCTTGGAGGCAATGTTTCTTAAACAGTTACATTTGCAAGTGTGTGGAAGGAAT 2400
Qy 2401 TGTTCGCGCAGCAGGTGACAGACCTGGAATCCCACTGTGGCATGGCAGTTGGTATA 2460
Db 2401 TGTTCGCGCAGCAGGTGACAGACCTGGAATCCCACTGTGGCATGGCAGTTGGTATA 2460
Qy 2461 CTGCTGACCAACCTTCTGTGATGTTGATTAATTTAGCAGTTGTGTTATCCGATTAAG 2520
Db 2461 CTGCTGACCAACCTTCTGTGATGTTGATTAATTTAGCAGTTGTGTTATCCGATTAAG 2520
Qy 2521 AAGGATTAAGGCAAGATTAATGTTAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Db 2521 AAGGATTAAGGCAAGATTAATGTTAAAGTCTCAAGCATCTGAAGTCAAACTCTGAGA 2580
Qy 2581 AGCTGAATTTGAAAGGAATGTTGAATTTATPAGCAATGTCATTTGCAACAACCA 2640
Db 2581 AGCTGAATTTGAAAGGAATGTTGAATTTATPAGCAATGTCATTTGCAACAACCA 2640
Qy 2641 TCTCATCTATTAATCTTTCATCTAACGTGATTAATTTTAAACAGATATTTCCCTCT 2700
Db 2641 TCTCATCTATTAATCTTTCATCTAACGTGATTAATTTTAAACAGATATTTCCCTCT 2700
Qy 2701 TGTCTTTAAATTTTGTGAATAATTTCTTTTGTAGGTGAGTCTGTGTCGCCAG 2760
Db 2701 TGTCTTTAAATTTTGTGAATAATTTCTTTTGTAGGTGAGTCTGTGTCGCCAG 2760

QY 2761 GCTGAGTACAGTGTGTGATCCAGCTCACTGACACTCCGCTCCGAGTTTCACATGA 2820
Db 2761 GCTGAGTACAGTGTGTGATCCAGCTCACTGACACTCCGCTCCGAGTTTCACATGA 2820
QY 2821 TTCTCCCTGCTCAGCTTCTTAAGTACTGGTTTACAGGCAACCAACCAATGCCAGCT 2880
Db 2821 TTCTCCCTGCTCAGCTTCTTAAGTACTGGTTTACAGGCAACCAACCAATGCCAGCT 2880
QY 2881 AATTTTGTATTTTAAATAGAGAGGGGTTTGGCAATTTGGCCAGGCTGGCTTGAATCTC 2940
Db 2881 AATTTTGTATTTTAAATAGAGAGGGGTTTGGCAATTTGGCCAGGCTGGCTTGAATCTC 2940
QY 2941 CTGACGTCAAGTATCTGCTGCTGCTTGTCTCCCAATACAGGCAATGACCACTGCAACCA 3000
Db 2941 CTGACGTCAAGTATCTGCTGCTGCTTGTCTCCCAATACAGGCAATGACCACTGCAACCA 3000
QY 3001 CCTACTTGAATATTTCAATGTGCTATAGACATTTAGAGATTTTCAATTTTCCATGACAT 3060
Db 3001 CCTACTTGAATATTTCAATGTGCTATAGACATTTAGAGATTTTCAATTTTCCATGACAT 3060
QY 3061 TTTTCTCTCTGCAAAATGGCTAGCTACTGTGTTTTCCTTTGGGGGCAAGACACT 3120
Db 3061 TTTTCTCTCTGCAAAATGGCTAGCTACTGTGTTTTCCTTTGGGGGCAAGACACT 3120
QY 3121 CATTAATATTTCTGATCATTTTCTTATCAAGAGATATATCATGTTGTCTCATAGA 3180
Db 3121 CATTAATATTTCTGATCATTTTCTTATCAAGAGATATATCATGTTGTCTCATAGA 3180
QY 3181 ACTGCTGATTTCTCATTTAATGTTTTTCTGATTTCTCATCTGTGTCTCCCTTATCTTAC 3240
Db 3181 ACTGCTGATTTCTCATTTAATGTTTTTCTGATTTCTCATCTGTGTCTCCCTTATCTTAC 3240
QY 3241 TCCCTTGTATTTCACTGATTTTCAAACTTTGTCAAGAGAGAGAGAGAGACTCAG 3300
Db 3241 TCCCTTGTATTTCACTGATTTTCAAACTTTGTCAAGAGAGAGAGAGAGACTCAG 3300
QY 3301 GAAAAATTAATTAATAAAGAACAGCCTTTTGGCGCGCGCAATTC 3345
Db 3301 GAAAAATTAATTAATAAAGAACAGCCTTTTGGCGCGCGCAATTC 3345

RESULT 3
US-09-962-436-302
; Sequence 302, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppel, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 302
; LENGTH: 3345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-302

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCTCTGACCACTGAATGAGAGAAAGAGACTTTTAAACCATTTTGTGACTTA 60
Db 1 GAATTCGCTCTGACCACTGAATGAGAGAAAGAGACTTTTAAACCATTTTGTGACTTA 60

QY 61 CAGAAAGAAATTTGAATTAAGAAACATATGATCTTCAGGCCATCTTCACTCCCTGTGT 120
Db 61 CAGAAAGAAATTTGAATTAAGAAACATATGATCTTCAGGCCATCTTCACTCCCTGTGT 120
QY 121 CTTCTTATGCTTTATTTTGGCACTGATATATGCGCAAGAGGGGAAGTTTATGACCCCTG 180
Db 121 CTTCTTATGCTTTATTTTGGCACTGATATATGCGCAAGAGGGGAAGTTTATGACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATGAAGGCAAGAACCCGATCAATTAATATTCAGATT 240
Db 181 AAACCCATGACATTTTCTATTTATGAAGGCAAGAACCCGATCAATTAATATTCAGATT 240
QY 241 AAGGCCAATCTCTGCTGTGACTTTTGAATCTAATCTGGGAGACAGACAAATATTTGTG 300
Db 241 AAGGCCAATCTCTGCTGTGACTTTTGAATCTAATCTGGGAGACAGACAAATATTTGTG 300
QY 301 ATGGAACGGAGGAGACTTCTGTATTAACACAGAGCTTTGACAGGGAACAAATCTACT 360
Db 301 ATGGAACGGAGGAGACTTCTGTATTAACACAGAGCTTTGACAGGGAACAAATCTACT 360
QY 361 CACAATCTCCAGGTTGACGCCCTGAGCGCTAATGGAATTAATGAGAGGGTCCAGTCCCT 420
Db 361 CACAATCTCCAGGTTGACGCCCTGAGCGCTAATGGAATTAATGAGAGGGTCCAGTCCCT 420
QY 421 ATCACAATGAAGTGAAGGACATCAACGACATGACCCACCTTCTCAGTCAAGATAC 480
Db 421 ATCACAATGAAGTGAAGGACATCAACGACATGACCCACCTTCTCAGTCAAGATAC 480
QY 481 GAAAGCTCAGTAAAGGAGAACTTGGCCGACAGAAAGCCCTTCTGTATGTCAATGCCACA 540
Db 481 GAAAGCTCAGTAAAGGAGAACTTGGCCGACAGAAAGCCCTTCTGTATGTCAATGCCACA 540
QY 541 GACCTGATGATATCCGCGCACTCCCAATGGCCACTTATTAACAAGATGTCAATCCAGCTT 600
Db 541 GACCTGATGATATCCGCGCACTCCCAATGGCCACTTATTAACAAGATGTCAATCCAGCTT 600
QY 601 CCCATGATCAACATGATCATGATCTTCAATCAACAAACGAGAGCATCTCTCT 660
Db 601 CCCATGATCAACATGATCATGATCTTCAATCAACAAACGAGAGCATCTCTCT 660
QY 661 ACCGAGAGGAGATCTCAGAAATGAAATCTGCTAAGAAATCTTCTATTAATCTGTGATC 720
Db 661 ACCGAGAGGAGATCTCAGAAATGAAATCTGCTAAGAAATCTTCTATTAATCTGTGATC 720
QY 721 TCAGTAAGGACATGGGAGGCGCAGATGGAATTCCTTCAAGATATCCACATCTGTGAT 780
Db 721 TCAGTAAGGACATGGGAGGCGCAGATGGAATTCCTTCAAGATATCCACATCTGTGAT 780
QY 781 ATCATAGTACAGAGATATTTTGAAGACCAAAACCTGTGAGATGTGAGAAATCTCA 840
Db 781 ATCATAGTACAGAGATATTTTGAAGACCAAAACCTGTGAGATGTGAGAAATCTCA 840
QY 841 ACTGATCTTCAACCCATCAAAATCACTCAGGTGCGGTGAATGATCCGGTGCACAAATAT 900
Db 841 ACTGATCTTCAACCCATCAAAATCACTCAGGTGCGGTGAATGATCCGGTGCACAAATAT 900
QY 901 TCCCTTGTGACAAAGAGAGCTGCCAAGATTCCTTCAATTTCAATTAACAAGAGAGAT 960
Db 901 TCCCTTGTGACAAAGAGAGCTGCCAAGATTCCTTCAATTTCAATTAACAAGAGAGAT 960
QY 961 ATTTACGTACTCAGCCCTTGTGACCGAGAGAAAGATGATATTTTATGACATT 1020
Db 961 ATTTACGTACTCAGCCCTTGTGACCGAGAGAAAGATGATATTTTATGACATT 1020
QY 1021 GCAAGAGATGAGTACGAGAAACCACTTCAATATCCCTGGAATTCATGTAAAGTTAA 1080
Db 1021 GCAAGAGATGAGTACGAGAAACCACTTCAATATCCCTGGAATTCATGTAAAGTTAA 1080
QY 1081 GATATTAATGATATATCCATCATGTCCGTACCGATATTTTATGAGTCCAGAG 1140
Db 1081 GATATTAATGATATATCCATCATGTCCGTACCGATATTTTATGAGTCCAGAG 1140
QY 1141 AATGAACGACTGGGTAAAGATATCGGACCTTTATGCAATGACAGGATGAAGAAAT 1200

Db	1141	AATGAAAGACTGGGGTAACAGTATGCGGACCCTTATCTGACATGACAGGGATGAAGAAAT	1200
Oy	1201	ACTGCAACAGATTTCCTTAACCTACAGAGTTGTGAGCAAACTCCCAACTTCCATGAT	1260
Db	1201	ACTGCCAACAGTTTCTTAACCTACAGAGTTGTGAGCAAACTCCCAACTTCCATGAT	1260
Oy	1261	GGACTCTTCTTAATCCAAACCTATGCTGGAAATGTACAGTTAGCTAAACAGTCTTGAAG	1320
Db	1261	GGACTCTTCTTAATCCAAACCTATGCTGGAAATGTACAGTTAGCTAAACAGTCTTGAAG	1320
Oy	1321	AAGCAAGTACTCTTCAGTACAACTTAAACGATTAAGGTGCTGACAAAGATTCAAAGC	1380
Db	1321	AAGCAAGTACTCTTCAGTACAACTTAAACGATTAAGGTGCTGACAAAGATTCAAAGC	1380
Oy	1381	CTTTGTTTGTGCAAAATCAACGTTAATGATATCATGATGATGATCCCATCTTTGAAGAA	1440
Db	1381	CTTTGTTTGTGCAAAATCAACGTTAATGATATCATGATGATGATCCCATCTTTGAAGAA	1440
Oy	1441	TCAGATTATGGAACCTGACTCTTGTGCTGAAGACAAACATTTGGTCCACCATCTTAAC	1500
Db	1441	TCAGATTATGGAACCTGACTCTTGTGCTGAAGACAAACATTTGGTCCACCATCTTAAC	1500
Oy	1501	ATCCAGGCCACTGATGCTGATGAGCCATTACTGGAGTTCTAAATTCGTATCATATC	1560
Db	1501	ATCCAGGCCACTGATGCTGATGAGCCATTACTGGAGTTCTAAATTCGTATCATATC	1560
Oy	1561	ATTAAGGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCCATACCAACCGGA	1620
Db	1561	ATTAAGGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCCATACCAACCGGA	1620
Oy	1621	TATGTCATTAATTAAGCCCTCTTGATTTTGAAACAGCAGCTGTTCCAACTTTGTTC	1680
Db	1621	TATGTCATTAATTAAGCCCTCTTGATTTTGAAACAGCAGCTGTTCCAACTTTGTTC	1680
Oy	1681	AAAGCAAAAAATCCTGAGCCTCTAGTGTGTTGTGTGAAGTACATGCAAGTCTTTTGCC	1740
Db	1681	AAAGCAAAAAATCCTGAGCCTCTAGTGTGTTGTGTGAAGTACATGCAAGTCTTTTGCC	1740
Oy	1741	AAGTTCACGCTTAATGTGACAGATGTGAATGAAGCACTCAATTTTCCCAACAGTATTC	1800
Db	1741	AAGTTCACGCTTAATGTGACAGATGTGAATGAAGCACTCAATTTTCCCAACAGTATTC	1800
Oy	1801	CAAGGAAAGTCAGTGAAGATGTAGCTATAGGCACTAAAGTGGGCAATGTAGTCCAG	1860
Db	1801	CAAGGAAAGTCAGTGAAGATGTAGCTATAGGCACTAAAGTGGGCAATGTAGTCCAG	1860
Oy	1861	GATCCAGAAAGTCTGAGACATTAAGCTATTCACTGAGGGAGACACAAGAGTTGGCTTAA	1920
Db	1861	GATCCAGAAAGTCTGAGACATTAAGCTATTCACTGAGGGAGACACAAGAGTTGGCTTAA	1920
Oy	1921	ATTGACCAAGTCACTGAGATGTGAGATCTTTAGTGTGGCTCCATTGGACAGAGAGCGGAAGT	1980
Db	1921	ATTGACCAAGTCACTGAGATGTGAGATCTTTAGTGTGGCTCCATTGGACAGAGAGCGGAAGT	1980
Oy	1981	CCATATCGGGTACAAAGTGTGGCCACAGAAGTGGGGGCTTCTTAAGCTCTGTGCA	2040
Db	1981	CCATATCGGGTACAAAGTGTGGCCACAGAAGTGGGGGCTTCTTAAGCTCTGTGCA	2040
Oy	2041	GAGTTCCACCTGATCTTATGATGTGAATGACAACTTCCAGGCTTACCCAAAGACTAC	2100
Db	2041	GAGTTCCACCTGATCTTATGATGTGAATGACAACTTCCAGGCTTACCCAAAGACTAC	2100
Oy	2101	ACGGGCTGTCTTCTGCAATGCCCTGAGTACCTGGAAGTCTCATTTTGAGGCTACT	2160
Db	2101	ACGGGCTGTCTTCTGCAATGCCCTGAGTACCTGGAAGTCTCATTTTGAGGCTACT	2160
Oy	2161	GATGATGATCAGACTTATTTTGGGGGTCCCAATTTTACATTTTCCCTGCGAGTGGAGC	2220
Db	2161	GATGATGATCAGACTTATTTTGGGGGTCCCAATTTTACATTTTCCCTGCGAGTGGAGC	2220
Oy	2221	TTACAAAAGCACTGGGAAGTTTCCAAAATCAATGCTACTCATGCCGACTGTCTTACAGG	2280
Db	2221	TTACAAAAGCACTGGGAAGTTTCCAAAATCAATGCTACTCATGCCGACTGTCTTACAGG	2280

Db	2221	TTACAAAACGACTGGGAAGTTTCCAAATCAATGTAAGTACTATGCCGACCTGTACAGG	2280
Qy	2281	CACA CAGACTTTGAGAGAGAGGGCTATGTCGTCCTTGATCCGCATCAATGATGGGGCTCGG	2340
Db	2281	CACA CAGACTTTGAGAGAGAGGGCTATGTCGTCCTTGATCCGCATCAATGATGGGGCTCGG	2340
Qy	2341	CCACCCTTGGAAAGGCATGTTTCTTTCACAGTTACATTTCCGACGTTGTGTGGAAGGAAT	2400
Db	2341	CCACCCTTGGAAAGGCATGTTTCTTTCACAGTTACATTTCCGACGTTGTGTGGAAGGAAT	2400
Qy	2401	TGTTTCCGGCAGAGAGTTCACACAGCTGGGATACCACTGTGGGCATGGGCAAGTGTGTATA	2460
Db	2401	TGTTTCCGGCAGAGAGTTCACACAGCTGGGATACCACTGTGGGCATGGGCAAGTGTGTATA	2460
Qy	2461	CTGCTGACCACTCTTGTGTGATGTTGTATAATTTTACAGCTGTGTATTCCGATTAAG	2520
Db	2461	CTGCTGACCACTCTTGTGTGATGTTGTATAATTTTACAGCTGTGTATTCCGATTAAG	2520
Qy	2521	AAGGATTAAGGCAGAAATATGTTGAAAGGCTCAACATCTGAAGTCAAAACCTCTGAGA	2580
Db	2521	AAGGATTAAGGCAGAAATATGTTGAAAGGCTCAACATCTGAAGTCAAAACCTCTGAGA	2580
Qy	2581	AGCTGAATTTGAAAGAATGTTGAAATTTATATAGCAAGTGCATTTTCAGCAACAACA	2640
Db	2581	AGCTGAATTTGAAAGAATGTTGAAATTTATATAGCAAGTGCATTTTCAGCAACAACA	2640
Qy	2641	TCTGATCTTAATTAATCTTTTCATCTAAGCTGCATTAATTTTAAACAGATATTCCTCT	2700
Db	2641	TCTGATCTTAATTAATCTTTTCATCTAAGCTGCATTAATTTTAAACAGATATTCCTCT	2700
Qy	2701	TGTCCTTAATATTTTGTCTAAATATTTCTTTTGAAGTGAAGTCTCTGTGCCAG	2760
Db	2701	TGTCCTTAATATTTTGTCTAAATATTTCTTTTGAAGTGAAGTCTCTGTGCCAG	2760
Qy	2761	GCTGAGATGACGTGTGTGATCCAGCTCACTGACACTCCGCTCTGTGGTTCACTAGA	2820
Db	2761	GCTGAGATGACGTGTGTGATCCAGCTCACTGACACTCCGCTCTGTGGTTCACTAGA	2820
Qy	2821	TTCTCTCGCTCAGACTTCTTAAGTAGCTGGGTTTACAGGCAACCAACATATGCCAGCT	2880
Db	2821	TTCTCTCGCTCAGACTTCTTAAGTAGCTGGGTTTACAGGCAACCAACATATGCCAGCT	2880
Qy	2881	AATTTTGTATTTTATATAGAGAGGGGTTTCGCATTTTGGCCAGGCTGTGTGAATC	2940
Db	2881	AATTTTGTATTTTATATAGAGAGGGGTTTCGCATTTTGGCCAGGCTGTGTGAATC	2940
Qy	2941	CTGACGTCAAGATGATCTGCCCTGCTGGTCTCCCAATACAGGCAATGAACCACTGCCA	3000
Db	2941	CTGACGTCAAGATGATCTGCCCTGCTGGTCTCCCAATACAGGCAATGAACCACTGCCA	3000
Qy	3001	CCTACTTAGATATTTCAATGTGCTATAGACATTTAGAGAGATTTTCAATTTTTCATGACAT	3060
Db	3001	CCTACTTAGATATTTCAATGTGCTATAGACATTTAGAGAGATTTTCAATTTTTCATGACAT	3060
Qy	3061	TTTTCTCTCTGCAATATGCTTACACTTGTGTTTTTCCCTTTTGGGGCAAGCAACT	3120
Db	3061	TTTTCTCTCTGCAATATGCTTACACTTGTGTTTTTCCCTTTTGGGGCAAGCAACT	3120
Qy	3121	CATTAAATATCTGACATTTTCTTTATCAAGAGATATATCAGTGTGTCTCATAGA	3180
Db	3121	CATTAAATATCTGACATTTTCTTTATCAAGAGATATATCAGTGTGTCTCATAGA	3180
Qy	3181	ACTGCGCTGATTCATTTATGTTTTTTCTGATTCATCTCTGTCTCCCTTCATCTTGAC	3240
Db	3181	ACTGCGCTGATTCATTTATGTTTTTTCTGATTCATCTCTGTCTCCCTTCATCTTGAC	3240
Qy	3241	TCCTTTGTAATTCATGAATTTCAACATTTGTCAAGAGAAAGAAAGTATGAGACTCAG	3300
Db	3241	TCCTTTGTAATTCATGAATTTCAACATTTGTCAAGAGAAAGAAAGTATGAGACTCAG	3300
Qy	3301	GAATAATTAATTAATTAAGAACAGCTTTTGGGGCCGCCAATTC	3345
Db	3301	GAATAATTAATTAATTAAGAACAGCTTTTGGGGCCGCCAATTC	3345

RESULT 4
US-09-833-263-1076
Sequence 1076, Application US/09833263
Patent No. US20020110547A1
GENERAL INFORMATION:
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Stolk, John A.
APPLICANT: Meagher, Madeleine J.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C12
CURRENT APPLICATION NUMBER: US/09/833,263
CURRENT FILING DATE: 2001-04-10
NUMBER OF SEQ ID NOS: 1093
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 1076
LENGTH: 3345
TYPE: DNA
ORGANISM: Homo sapiens
US-09-833-263-1076

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGCTCCAGCCCTGATGAGAAAGGACCTTTTACCAACATTTTGTGACTTA 60
DB 1 GAATTCGCTCCAGCCCTGATGAGAAAGGACCTTTTACCAACATTTTGTGACTTA 60
QY 61 CAGAAAGGAAATTTGAATTAAGAAAACATATGATACCTTCAAGGCCATCTTCACTCCCTGT 120
DB 61 CAGAAAGGAAATTTGAATTAAGAAAACATATGATACCTTCAAGGCCATCTTCACTCCCTGT 120
QY 121 CTTCTTATGCTTTATTTGGCACTGATATGATGAGGAGGAGGAAATTTAGTGAACCCCTG 180
DB 121 CTTCTTATGCTTTATTTGGCACTGATATGATGAGGAGGAGGAAATTTAGTGAACCCCTG 180
QY 181 AAACCCATGACATTTTCTATTTATTAAGAGCCAGAACCCAGATCAAAATTAATTCAGATT 240
DB 181 AAACCCATGACATTTTCTATTTATTAAGAGCCAGAACCCAGATCAAAATTAATTCAGATT 240
QY 241 AAGGCCAATCTCTGCTGATGACTTTTGAACCTAAGTGGGAGACAGACATATTTGTG 300
DB 241 AAGGCCAATCTCTGCTGATGACTTTTGAACCTAAGTGGGAGACAGACATATTTGTG 300
QY 301 ATAGAAAGGAGGAGCTTCTGTATTAACAAGACCTTGAACAGGAAACAAGATCTACT 360
DB 301 ATAGAAAGGAGGAGCTTCTGTATTAACAAGACCTTGAACAGGAAACAAGATCTACT 360
QY 361 CACAAATCTCCAGGTTGGAGCCCTGAGAGCTTAATGAAATTAATGAGGAGCTCAATCCCT 420
DB 361 CACAAATCTCCAGGTTGGAGCCCTGAGAGCTTAATGAAATTAATGAGGAGCTCAATCCCT 420
QY 421 ATCAACATAGAGTGAAGGACATCAACAGACATGACCCAGTTTCCAGTCAAGATAC 480
DB 421 ATCAACATAGAGTGAAGGACATCAACAGACATGACCCAGTTTCCAGTCAAGATAC 480
QY 481 GAAAGGCTCAGTAAAGGAGAACTCTCGCCAGAGAAAGCCCTTCTGTATGTCATATGCACA 540
DB 481 GAAAGGCTCAGTAAAGGAGAACTCTCGCCAGAGAAAGCCCTTCTGTATGTCATATGCACA 540
QY 541 GACCTGATGATATCCGGCCATCTCCCAATGAGCAGCTTTATTAACAAGATGTCATCAAGCTT 600
DB 541 GACCTGATGATATCCGGCCATCTCCCAATGAGCAGCTTTATTAACAAGATGTCATCAAGCTT 600
QY 601 CCCATGATCAACAATGTCATGTAATCTTTCAATCAACAACAAAGGAGCAGATCTCTCT 660
DB 601 CCCATGATCAACAATGTCATGTAATCTTTCAATCAACAACAAAGGAGCAGATCTCTCT 660
QY 661 ACCGAGAGGAGATCTGAGAAATTCCTGCTAAGATCTCTTCTATATCTGTGATC 720

DB 661 ACCGAGAGGAGATCTGAGAAATTCCTGCTAAGATCTCTTCTATATCTGTGATC 720
QY 721 TCAGTGAAGGACATGGAGGAGCCAGAGTGAATTCCTTCAAGTATACCAATCTGTGAT 780
DB 721 TCAGTGAAGGACATGGAGGAGCCAGAGTGAATTCCTTCAAGTATACCAATCTGTGAT 780
QY 781 ATCATAGTGAACAAGATATTTGAAAGACACCAAAACCTGTGAGATGTGAGAAACTCA 840
DB 781 ATCATAGTGAACAAGATATTTGAAAGACACCAAAACCTGTGAGATGTGAGAAACTCA 840
QY 841 ACTGATCTCAACCCCATCAAAATCACTCAAGTCCGCTGAGATATCCCGGTGACATAT 900
DB 841 ACTGATCTCAACCCCATCAAAATCACTCAAGTCCGCTGAGATATCCCGGTGACATAT 900
QY 901 TCCTTATGTTGAACAAGAGAGGCTGCAAGATTCCTTCAATTTGACATGACAGAGAT 960
DB 901 TCCTTATGTTGAACAAGAGAGGCTGCAAGATTCCTTCAATTTGACATGACAGAGAT 960
QY 961 ATTTACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 ATTTACGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 GCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 GATATTAATGATATATCACTCAATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 GATATTAATGATATATCACTCAATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 AATGAAGCACTGGGTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 AATGAAGCACTGGGTAAAGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ACTGCAACAGATTTTCTAAACCTAAGATGATGATGATGATGATGATGATGATGAT 1260
DB 1201 ACTGCAACAGATTTTCTAAACCTAAGATGATGATGATGATGATGATGATGATGAT 1260
QY 1261 GGAATCTTCTTAATCCAAACCTATGCTGATGATGATGATGATGATGATGATGAT 1320
DB 1261 GGAATCTTCTTAATCCAAACCTATGCTGATGATGATGATGATGATGATGATGAT 1320
QY 1321 AAGCAAGATATCTCAGTAACTTAAGATGATGATGATGATGATGATGATGATGAT 1380
DB 1321 AAGCAAGATATCTCAGTAACTTAAGATGATGATGATGATGATGATGATGATGAT 1380
QY 1381 CTTTGTGTTGTGCAATCAACCTTATGATATCAATGATGATGATGATGATGATGAT 1440
DB 1381 CTTTGTGTTGTGCAATCAACCTTATGATATCAATGATGATGATGATGATGATGAT 1440
QY 1441 TCAGATTAAGGAACTGATCTTCTGTAAGACAAACATTTGGGTCACATCTTAACC 1500
DB 1441 TCAGATTAAGGAACTGATCTTCTGTAAGACAAACATTTGGGTCACATCTTAACC 1500
QY 1501 ATCCAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1501 ATCCAGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
QY 1561 ATTAAGGAGAGAGTGAAGGAGGCTGAGGAGTGAACAGATGCCATACCAACCGGA 1620
DB 1561 ATTAAGGAGAGAGTGAAGGAGGCTGAGGAGTGAACAGATGCCATACCAACCGGA 1620
QY 1621 TATGTCATTAATTAAGAGCTTCTGTAAGACAGAGCTGTTTCAACATGTGTTCC 1680
DB 1621 TATGTCATTAATTAAGAGCTTCTGTAAGACAGAGCTGTTTCAACATGTGTTCC 1680
QY 1681 AAAGAGAGAAATCTGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 AAAGAGAGAAATCTGAGGCTGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 AAGTTCACGCTTATTTGTCAGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 AAGTTCACGCTTATTTGTCAGATGATGATGATGATGATGATGATGATGATGAT 1800

Db 1741 AAGTTACGCTTATTTGACAGATGATGATGAAGACCTCAATTTTCCAAACGATATTC 1800
Qy 1801 CAAGCGAAGTCACTGAGAGATGTAAGTGAAGCACTAAAGTGGCAATGTGACCTGCAAG 1860
Db 1801 CAAGCGAAGTCACTGAGAGATGTAAGTGAAGCACTAAAGTGGCAATGTGACCTGCAAG 1860
Qy 1861 GATCCGAAGTGTGAGCATTAAGCTATTCATGAGGGGAGACCAAGAGTGGCTTAA 1920
Db 1861 GATCCGAAGTGTGAGCATTAAGCTATTCATGAGGGGAGACCAAGAGTGGCTTAA 1920
Qy 1921 ATTGACCACTGACCTGAGATGATCTTTAGTGGCTCATTTGACAGAGAGCCGAGAT 1980
Db 1921 ATTGACCACTGACCTGAGATGATCTTTAGTGGCTCATTTGACAGAGAGCCGAGAT 1980
Qy 1981 CCATATCGGGTACAGTGTGGCCACAGAGTGAAGGGGGCTTCTTAAAGTCTGTCTCA 2040
Db 1981 CCATATCGGGTACAGTGTGGCCACAGAGTGAAGGGGGCTTCTTAAAGTCTGTCTCA 2040
Qy 2041 GAGTTCCACCTGATCTTATGATGATGAATGACAAACCTCCAGGCTAGCCAGAGATAC 2100
Db 2041 GAGTTCCACCTGATCTTATGATGATGAATGACAAACCTCCAGGCTAGCCAGAGATAC 2100
Qy 2101 ACGGGCTTGTCTCTGCAATCCCTGAGTCACTGAGAGTCTCATTTTTCGAGGCTACT 2160
Db 2101 ACGGGCTTGTCTCTGCAATCCCTGAGTCACTGAGAGTCTCATTTTTCGAGGCTACT 2160
Qy 2161 GATGATATGACACTTATTTGGGGTCCCATTTTACATTTTCCCTGCGAGTGGAAAC 2220
Db 2161 GATGATATGACACTTATTTGGGGTCCCATTTTACATTTTCCCTGCGAGTGGAAAC 2220
Qy 2221 TTTCAAAAGCATGGGAGATTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTTCAAAAGCATGGGAGATTTCCAAATCAATGATGATGATGATGATGATGATGATGAT 2280
Qy 2281 CACACAGACTTGAAGAGAGGCGATGTCGTCTTGAATCCGATCAATGATGATGATGATGAT 2340
Db 2281 CACACAGACTTGAAGAGAGGCGATGTCGTCTTGAATCCGATCAATGATGATGATGATGAT 2340
Qy 2341 CCACCTTGGAGAGATTTGTTTCTTACAGATTAATTTGCAAGTGTGTGTGAAGAAAT 2400
Db 2341 CCACCTTGGAGAGATTTGTTTCTTACAGATTAATTTGCAAGTGTGTGTGAAGAAAT 2400
Qy 2401 TGTTCGGCGAGAGGATGACACAGATGGGATCCACTGGGCGATGGAGTGTGATTA 2460
Db 2401 TGTTCGGCGAGAGGATGACACAGATGGGATCCACTGGGCGATGGAGTGTGATTA 2460
Qy 2461 CTGCTGACCACTTCTGTGATGATGATTAATTTTGAAGATGATGATGATGATGATGAT 2520
Db 2461 CTGCTGACCACTTCTGTGATGATGATTAATTTTGAAGATGATGATGATGATGATGAT 2520
Qy 2521 AAGGATTAAGGCAAGATATGTTGAAGTGTCTCAAGATGTAAGTCAAACTCTGAGA 2580
Db 2521 AAGGATTAAGGCAAGATATGTTGAAGTGTCTCAAGATGTAAGTCAAACTCTGAGA 2580
Qy 2581 AGCTGAATTTGAAGGATGTTGAATTAATTAATGCAAGTCTATTGACCAACCA 2640
Db 2581 AGCTGAATTTGAAGGATGTTGAATTAATTAATGCAAGTCTATTGACCAACCA 2640
Qy 2641 TCTGATCTTATTTACTTTTCACTAAGTCACTAATTAATTTTAAACAGATTTCCCTCT 2700
Db 2641 TCTGATCTTATTTACTTTTCACTAAGTCACTAATTAATTTTAAACAGATTTCCCTCT 2700
Qy 2701 TGTCTTATTAATTTGCTAAATTAATTTTGTGAGTGTGCTGTGCTGTGCTGTGCTGT 2760
Db 2701 TGTCTTATTAATTTGCTAAATTAATTTTGTGAGTGTGCTGTGCTGTGCTGTGCTGT 2760
Qy 2761 GCTGAGAGTCAAGTGTGATGATCCAGCTCACTGCACTCCGCTCTCTGAGGATCACTGA 2820
Db 2761 GCTGAGAGTCAAGTGTGATGATCCAGCTCACTGCACTCCGCTCTCTGAGGATCACTGA 2820
Qy 2821 TTTCTCTGCTCAGCTTCTTAAGTATGCTGGGTTTACAGGCAACCAACCACTGCTGCT 2880
Db 2821 TTTCTCTGCTCAGCTTCTTAAGTATGCTGGGTTTACAGGCAACCAACCACTGCTGCT 2880

Qy 2881 AATTTTGTATTTTATATAGAGACGGGGTTTGCCATTTGGCCAGGCTGTCTTGAATC 2940
Db 2881 AATTTTGTATTTTATATAGAGACGGGGTTTGCCATTTGGCCAGGCTGTCTTGAATC 2940
Qy 2941 CTGACGTCAAGTATGCTGCTGTGCTGTCTGCTCAATACAGGATGATGATGATGATGAT 3000
Db 2941 CTGACGTCAAGTATGCTGCTGTGCTGTGCTGTGCTCAATACAGGATGATGATGATGAT 3000
Qy 3001 CCTACTTAAGTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
Db 3001 CCTACTTAAGTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060
Qy 3061 TTTTCCCTCTGCAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Db 3061 TTTTCCCTCTGCAATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3120
Qy 3121 CATTAATATTTCTGATCAATTTTCTTATACAGAGATATATATATATATATATATATAT 3180
Db 3121 CATTAATATTTCTGATCAATTTTCTTATACAGAGATATATATATATATATATATATAT 3180
Qy 3181 ACTGCTGATTTCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 ACTGCTGATTTCAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Qy 3241 TCTTGTGATTTTCACTGATTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 3300
Db 3241 TCTTGTGATTTTCACTGATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 3300
Qy 3301 GAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3345
Db 3301 GAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3345

RESULT 5
US-09-880-107-3288
; Sequence 3288, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3288
; LENGTH: 3345
; TYPE: DNA
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U07969
US-09-880-107-3288

Query Match 100.0%; Score 3345; DB 10; Length 3345;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3345; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAATTCGCTCGACCACTGAATGAGAAAGAACTTTTAAACCAATTTGTGACTTA 60
Db 1 GAATTCGCTCGACCACTGAATGAGAAAGAAAGAACTTTTAAACCAATTTGTGACTTA 60
Qy 61 CAGAAAGAAATTTGAATTAAGAAATCTATATCTTCAAGGCCATCTTCACTCCCTGTGT 120
Db 61 CAGAAAGAAATTTGAATTAAGAAATCTATATCTTCAAGGCCATCTTCACTCCCTGTGT 120

QY 121 CTTCTTATGCTTATTTGGCACTGATATGCGCAAGAGGGAGTTAGTGAAGCCCTG 180
Db 121 CTTCTTATGCTTATTTGGCACTGATATGCGCAAGAGGGAGTTAGTGAAGCCCTG 180
QY 181 AAAGCCATGACATTTCTATTTATGAGGCGCAAGACCGAGTCAATTAATTTCCAGTTT 240
Db 181 AAAGCCATGACATTTCTATTTATGAGGCGCAAGACCGAGTCAATTAATTTCCAGTTT 240
QY 241 AAGGCCAATCCTCTGCTGTGATCTTTGAACTAACTGGGAGACAGCAATATTTGTG 300
Db 241 AAGGCCAATCCTCTGCTGTGATCTTTGAACTAACTGGGAGACAGCAATATTTGTG 300
QY 301 ATAGAACGGAGGAGACTTCTGTATTAACAAGACCTTTGACAGGAAACAAAGATCTACT 360
Db 301 ATAGAACGGAGGAGACTTCTGTATTAACAAGACCTTTGACAGGAAACAAAGATCTACT 360
QY 361 CACAAATCCAGGTTGAGAGCCCTGAGAGCTTAATGGAATTAATGAGAGGCTCCAGTCCCT 420
Db 361 CACAAATCCAGGTTGAGAGCCCTGAGAGCTTAATGGAATTAATGAGAGGCTCCAGTCCCT 420
QY 421 ATCAACATAGAGGAGACATCAACAGCAATGCAACCGTTTCTCCAGTCAAGATAC 480
Db 421 ATCAACATAGAGGAGACATCAACAGCAATGCAACCGTTTCTCCAGTCAAGATAC 480
QY 481 GAAAGCTCAGTAAAGGAGAACTCTGCTGCAAGAAAGCCCTTTCTGTATGTCAATGCCACA 540
Db 481 GAAAGCTCAGTAAAGGAGAACTCTGCTGCAAGAAAGCCCTTTCTGTATGTCAATGCCACA 540
QY 541 GACCTGATATATCCGGCACTCCCAATGGCCAGTTTATTAACAGATTTGTCATCCAGTT 600
Db 541 GACCTGATATATCCGGCACTCCCAATGGCCAGTTTATTAACAGATTTGTCATCCAGTT 600
QY 601 CCCATGATCAACATGTCATGTAATCTTCAATCAACAAACGAGGAGCATCTCTCTT 660
Db 601 CCCATGATCAACATGTCATGTAATCTTCAATCAACAAACGAGGAGCATCTCTCTT 660
QY 661 ACCCGAGAGGGATCTCAGAAATGGAATCTGCTAAGAAATCTTCTTAATTCGTGTATC 720
Db 661 ACCCGAGAGGGATCTCAGAAATGGAATCTGCTAAGAAATCTTCTTAATTCGTGTATC 720
QY 721 TCACTGAAAGACATGAGAGGCGCAGATGAGAAATTTCTTCACTGATATCAATCTGTGAT 780
Db 721 TCACTGAAAGACATGAGAGGCGCAGATGAGAAATTTCTTCACTGATATCAATCTGTGAT 780
QY 781 ATCATAGTGCAGAGAAATATTTGGAAGAGCAACAAACCTGTGAGATGAGTGAAGAAATCA 840
Db 781 ATCATAGTGCAGAGAAATATTTGGAAGAGCAACAAACCTGTGAGATGAGTGAAGAAATCA 840
QY 841 ACTGATCTCAGCCCATCAAAATCACTCAGGTGCGGTGAGATATCCCGGTGACATAT 900
Db 841 ACTGATCTCAGCCCATCAAAATCACTCAGGTGCGGTGAGATATCCCGGTGACATAT 900
QY 901 TCCCTTATGTCAGAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTGACAGAAAGAGAT 960
Db 901 TCCCTTATGTCAGAAAGAGAGCTGCGCAAGATTTCCATTTTCAATTGACAGAAAGAGAT 960
QY 961 ATTATGTCAGTCAAGCCCTTGAACCGAGAGAAAGATGCAATATGTTTTTATGACAT 1020
Db 961 ATTATGTCAGTCAAGCCCTTGAACCGAGAGAAAGATGCAATATGTTTTTATGACAT 1020
QY 1021 GCAAAAGATAGTACGAGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
Db 1021 GCAAAAGATAGTACGAGAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
QY 1081 GATATTAATGATATCACTACATATGTCGTACACAGTAACCGTATTTGAGTCAGAG 1140
Db 1081 GATATTAATGATATCACTACATATGTCGTACACAGTAACCGTATTTGAGTCAGAG 1140
QY 1141 AATGACGACTGGGTAAAGATTCGGGACCTTTACTGACATGACAGGAGTAAAGAAAT 1200
Db 1141 AATGACGACTGGGTAAAGATTCGGGACCTTTACTGACATGACAGGAGTAAAGAAAT 1200
QY 1201 ACTGCCAAAGTTTCTAACTACAGGATTTGTGAGCAAACTCCCAAATCTTCCATGAT 1260

Db 1201 ACTGCCAAAGTTTCTAACTACAGGATTTGTGAGCAAACTCCCAAATCTTCCATGAT 1260
QY 1261 GGAATCTTCTAATCCAAACCTATGCTGGAATGTTACAGTACTAAACGCTCTTAAG 1320
Db 1261 GGAATCTTCTAATCCAAACCTATGCTGGAATGTTACAGTACTAAACGCTCTTAAG 1320
QY 1321 AAGCAAGATACCTCCTCAGTACAACTTAACGATAGAGGTGCTGACAAAGATTTCAAGAC 1380
Db 1321 AAGCAAGATACCTCCTCAGTACAACTTAACGATAGAGGTGCTGACAAAGATTTCAAGAC 1380
QY 1381 CTTTGTGTTGTGCAAAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
Db 1381 CTTTGTGTTGTGCAAAATCAACGTTATGATATCAATGATCAGATCCCATCTTTGAAAAA 1440
QY 1441 TCAATTAATGGAACCTGACCTCTGCTGGAAGACAAACATTTGGGTCCACATCTTAACC 1500
Db 1441 TCAATTAATGGAACCTGACCTCTGCTGGAAGACAAACATTTGGGTCCACATCTTAACC 1500
QY 1501 ATCCAGGCCCATGATGCTGATGAGCCATTTAATGAGGAGTTCTAAATTTCTGTATCATATC 1560
Db 1501 ATCCAGGCCCATGATGCTGATGAGCCATTTAATGAGGAGTTCTAAATTTCTGTATCATATC 1560
QY 1561 ATAAAGGAGACAGTGAAGGAGCGCTGAGGAGTTGACACAGATCCCATACCAACCCGGA 1620
Db 1561 ATAAAGGAGACAGTGAAGGAGCGCTGAGGAGTTGACACAGATCCCATACCAACCCGGA 1620
QY 1621 TATGTCAATATTAATAAGCTCTTGAATTTGAAACAGCAGCTGTTTCCAACTGTGTC 1680
Db 1621 TATGTCAATATTAATAAGCTCTTGAATTTGAAACAGCAGCTGTTTCCAACTGTGTC 1680
QY 1681 AAAGCAGAAAAATCTGAGCCCTAGTGTGTTGAGTGAAGTGAACATGCAATGCTTTTGCC 1740
Db 1681 AAAGCAGAAAAATCTGAGCCCTAGTGTGTTGAGTGAAGTGAACATGCAATGCTTTTGCC 1740
QY 1741 AAGTTACCGTTATTTGTCAGATGTAATGAAGCACTCAATTTTCCCAACGATATC 1800
Db 1741 AAGTTACCGTTATTTGTCAGATGTAATGAAGCACTCAATTTTCCCAACGATATC 1800
QY 1801 CAAAGCAAGTCACTGAGATGTAAGTATGAGCACTTAAGTGGCAATGTATGCCAAG 1860
Db 1801 CAAAGCAAGTCACTGAGATGTAAGTATGAGCACTTAAGTGGCAATGTATGCCAAG 1860
QY 1861 GATCCAGAAAGTGTGACATTAAGCTATTCATGAGGAGGAGACAAAGAGGTGGCTTAA 1920
Db 1861 GATCCAGAAAGTGTGACATTAAGCTATTCATGAGGAGGAGACAAAGAGGTGGCTTAA 1920
QY 1921 ATTGACACGTCAGCTGATGATCTTTAGTGTGCTCAATTGACAGAGAGCCGGAAT 1980
Db 1921 ATTGACACGTCAGCTGATGATCTTTAGTGTGCTCAATTGACAGAGAGCCGGAAT 1980
QY 1981 CCATATCGGGTACAAATGCTGTGCGCACAGAAATGAGGGGGGCTTCTTAAAGCTGTGTCA 2040
Db 1981 CCATATCGGGTACAAATGCTGTGCGCACAGAAATGAGGGGGGCTTCTTAAAGCTGTGTCA 2040
QY 2041 GAGTCCACCTGATCTTATGATGTAATGTAAGCAACCTCCAGGCTAGCCAGAGCTAC 2100
Db 2041 GAGTCCACCTGATCTTATGATGTAATGTAAGCAACCTCCAGGCTAGCCAGAGCTAC 2100
QY 2101 AGGGGCTGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTAATTTTCAAGGCTACT 2160
Db 2101 AGGGGCTGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTAATTTTCAAGGCTACT 2160
QY 2161 GATGATGATCAGACATTAATTTGGGGGTCCCATTTTCAATTTTCCCTGGCAGTGAAGC 2220
Db 2161 GATGATGATCAGACATTAATTTGGGGGTCCCATTTTCAATTTTCCCTGGCAGTGAAGC 2220
QY 2221 TTAACAAACGACGAGGAGGTTTCCAAATATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 TTAACAAACGACGAGGAGGTTTCCAAATATGATGATGATGATGATGATGATGATGATGAT 2280
QY 2281 CACACAGACTTTGAGAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340

D	b	2281	CACACACACTTTGAGAGAGGGCGGTATGTGCTCTTAAATCCGATCAATGATGGGGGTCCG	2340
Q	y	2341	CCACCCTTGGAGGCACTTTGTTCTTTTACCACTTACATTTCTGCAGTTGTGTGAAAGAACT	2400
D	b	2341	CCACCCTTGGAGGCACTTTGTTCTTTTACCACTTACATTTCTGCAGTTGTGTGAAAGAACT	2400
Q	y	2401	TGTTCCGGCCAGAGGCAACCACTGGGATACCACTGGGGCATGGCGAGTTGGATATA	2460
D	b	2401	TGTTCCGGCCAGAGGCAACCACTGGGATACCACTGGGGCATGGCGAGTTGGATATA	2460
Q	y	2461	CTGCTGACCAACCTTCTGTGTGATTTGATATATTTTAAAGCATTTGTATTCCGATTAAG	2520
D	b	2461	CTGCTGACCAACCTTCTGTGTGATTTGATATATTTTAAAGCATTTGTATTCCGATTAAG	2520
Q	y	2521	AAGGATAAAGGCAAAAGATATGTGTAAGTGTCTCAAGCATCTGAAGTCAAACTCTAGA	2580
D	b	2521	AAGGATAAAGGCAAAAGATATGTGTAAGTGTCTCAAGCATCTGAAGTCAAACTCTAGA	2580
Q	y	2581	AGCTGAATTTGAAAAGGAAAGTTTGAATTTATATAGCAAGGCAATTTCCAGCAACA	2640
D	b	2581	AGCTGAATTTGAAAAGGAAAGTTTGAATTTATATAGCAAGGCAATTTCCAGCAACA	2640
Q	y	2641	TCTCATCTTATCTTATCTTTCATCTCAACGCGCATTAATTTTAAACAGATATTCCTCT	2700
D	b	2641	TCTCATCTTATCTTATCTTTCATCTCAACGCGCATTAATTTTAAACAGATATTCCTCT	2700
Q	y	2701	TGTCCTTTAATATTTTGTCTAAATATTTCTTTTGTAGTGAAGTCTGTGTGCCAG	2760
D	b	2701	TGTCCTTTAATATTTTGTCTAAATATTTCTTTTGTAGTGAAGTCTGTGTGCCAG	2760
Q	y	2761	GCTGAGATACAGTGTGTGATCCCAAGCTCACTGAACCTCGGCGCTCCGGGTTCAACATA	2820
D	b	2761	GCTGAGATACAGTGTGTGATCCCAAGCTCACTGAACCTCGGCGCTCCGGGTTCAACATA	2820
Q	y	2821	TTCTCCCTCCAGCTTCTCTAAGTAGAGTGGGTTTACAGGCAACCAACCAATGCCAGCT	2880
D	b	2821	TTCTCCCTCCAGCTTCTCTAAGTAGAGTGGGTTTACAGGCAACCAACCAATGCCAGCT	2880
Q	y	2881	AATTTTGTATTTTAAATAGAGACGGGGTTTTGCGCAATTTGGCCAGGCTGTCTTGAATCT	2940
D	b	2881	AATTTTGTATTTTAAATAGAGACGGGGTTTTGCGCAATTTGGCCAGGCTGTCTTGAATCT	2940
Q	y	2941	CTGAGCTCAAGTGAATCTGCTGCTTGGTCTCCCAATACAGGCATGAACCATGCAACCA	3000
D	b	2941	CTGAGCTCAAGTGAATCTGCTGCTTGGTCTCCCAATACAGGCATGAACCATGCAACCA	3000
Q	y	3001	CCCTACTGATATTTTCATGTGCTATAGACATTTAGAGAGATTTTTCATTTTCCATGACAT	3060
D	b	3001	CCCTACTGATATTTTCATGTGCTATAGACATTTAGAGAGATTTTTCATTTTCCATGACAT	3060
Q	y	3061	TTTTCTCTCTGCAAAATGGCTTAGCTACTGTGTTTTCCCTTTTGGGGCAAGACAGACT	3120
D	b	3061	TTTTCTCTCTGCAAAATGGCTTAGCTACTGTGTTTTCCCTTTTGGGGCAAGACAGACT	3120
Q	y	3121	CATTAAATATCTGTACATTTTTCTTTATCAAGAGATATATCTAGTGTGTCTCATAGA	3180
D	b	3121	CATTAAATATCTGTACATTTTTCTTTATCAAGAGATATATCTAGTGTGTCTCATAGA	3180
Q	y	3181	ACTGCTGAGATTCATTTATGTTTTTTCGATTCATCTGCTGTCCTTCACTCTTAC	3240
D	b	3181	ACTGCTGAGATTCATTTATGTTTTTTCGATTCATCTGCTGTCCTTCACTCTTAC	3240
Q	y	3241	TCCTTTGGTATTTCACTGATTTTCAACAATTTGTCCAGAGAGAAAAAGTGAAGCATGAG	3300
D	b	3241	TCCTTTGGTATTTCACTGATTTTCAACAATTTGTCCAGAGAGAAAAAGTGAAGCATGAG	3300
Q	y	3301	GAATAATTAATTAATTAAGAAACAGCCTTTTGGGCGCGGAATTC	3345
D	b	3301	GAATAATTAATTAATTAAGAAACAGCCTTTTGGGCGCGGAATTC	3345

```

; Sequence 118 Application US/03962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Eppert, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0.
; SEQ ID NO 118
; LENGTH: 3654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-118

Query Match          99.1%; Score 3314.8; DB 10; Length 3654;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

Query Match	Similarity	99.1%	Score 3314.8	DB 10	Length 3654
Best Local	Similarity	99.8%	Pred. No. 0		
Matches 3319	Conservative	0	Mismatches	7	Indels 0
					Gaps 0
QY	8	GTCTGACCACTGATGGAGAGAAAGACCTTTAAACCAACATTTGTGACTTACAGAAAG	67		
DB	13	GTCTGACCACTGATGGAGAGAAAGACCTTTAAACCAACATTTGTGACTTACAGAAAG	72		
QY	68	GAATTGTAATTAAGAAACTATGATCTTCAGGGCCATCTTCACTCCCTGTGCTTCTTA	127		
DB	73	GAATTGTAATTAAGAAACTATGATCTTCAGGGCCATCTTCACTCCCTGTGCTTCTTA	132		
QY	128	TGCTTTATTTGGCAACTGATATGAGCCCAAGGGGAAAGTTAGTGAACCCCTGAAACCCA	187		
DB	133	TGCTTTATTTGGCAACTGATATGAGCCCAAGGGGAAAGTTAGTGAACCCCTGAAACCCA	192		
QY	188	TGACATTTTCTAATTTATGAGAGCCCAAGACCGAGTCAATATATATCCAGTTTAAAGCCA	247		
DB	193	TGACATTTTCTAATTTATGAGAGCCCAAGACCGAGTCAATATATATCCAGTTTAAAGCCA	252		
QY	248	ATCTCTCTGCTGTGACTTTTGAATCTAGGGGAGACAGACAATAATTTGTGATAGAAC	307		
DB	253	ATCTCTCTGCTGTGACTTTTGAATCTAGGGGAGACAGACAATAATTTGTGATAGAAC	312		
QY	308	GGGAGGGAGCTTCTGATATTAACAAGAGCCTTGGACAGGGAAACAAGATCTACTCAATC	367		
DB	313	GGGAGGGAGCTTCTGATATTAACAAGAGCCTTGGACAGGGAAACAAGATCTACTCAATC	372		
QY	368	TCCAGGTTGACGCCCTGAGCGCTATATGAGAGGTTCCAGTCCCTATCACCA	427		
DB	373	TCCAGGTTGACGCCCTGAGCGCTATATGAGAGGTTCCAGTCCCTATCACCA	432		
QY	428	TAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCAGTCAAGTACGAAGGCT	487		
DB	433	TAGAAGTGAAGACATCAACGACATGACCCAGTTTCTCAGTCAAGTACGAAGGCT	492		
QY	488	CAGTAGGACGAAACCTCGCCACGAGAAAGCCCTTTGTATGTCAATCCACAGACTGG	547		
DB	493	CAGTAGGACGAAACCTCGCCACGAGAAAGCCCTTTGTATGTCAATCCACAGACTGG	552		
QY	548	ATGATCCGGGCACTCCCAATGGCCAGCTTTATTAACAAGATTGTCAATCCAGTTCCTATGA	607		
DB	553	ATGATCCGGGCACTCCCAATGGCCAGCTTTATTAACAAGATTGTCAATCCAGTTCCTATGA	612		
QY	608	TCAACAATGTCAATGATCTTCAATCAACAACAAGGAGGCCATCTCTCTTAACCCAG	667		
DB	613	TCAACAATGTCAATGATCTTCAATCAACAACAAGGAGGCCATCTCTCTTAACCCAG	672		
QY	668	AGGAGATTCAGGAATTGAATCTGTGTAAAGATCCTTCTAATATGTGATGATCTCAGTGA	727		
DB	673	AGGAGATTCAGGAATTGAATCTGTGTAAAGATCCTTCTAATATGTGATGATCTCAGTGA	732		

QY 728 AGGACATGGAGGCGGCAAGTGAATTCCTTCACTGATACCAATCTGTGATATCATAG 787
DB 733 AGGACATGGAGGCGGCAAGTGAATTCCTTCACTGATACCAATCTGTGATATCATAG 792
QY 788 TGACAGAGATATTTGGAAAGCACAAACCTGTGAGATGTGTGAGAAAACCTCACTATC 847
DB 793 TGACAGAGATATTTGGAAAGCACAAACCTGTGAGATGTGTGAGAAAACCTCACTATC 852
QY 848 CTCACCCCATCAAAATCACTCAGGTGGGTGAATGATCCCGGTGACAAATATCTTAG 907
DB 853 CTCACCCCATCAAAATCACTCAGGTGGGTGAATGATCCCGGTGACAAATATCTTAG 912
QY 908 TTGCAAAAGAGAGCTGCGCAAGATTCCTTCAATTTGACACAGAAAGAGATATTTACG 967
DB 913 TTGCAAAAGAGAGCTGCGCAAGATTCCTTCAATTTGACACAGAAAGAGATATTTACG 972
QY 968 TGACTCAGCCCTTGGACCGGAGAAAGATGATATGTTTTTATGACGTTGCAAGG 1027
DB 973 TGACTCAGCCCTTGGACCGGAGAAAGATGATATGTTTTTATGACGTTGCAAGG 1032
QY 1028 ATGAGTACGGAACCACTTTCAATATCCGTGAAATTTGATTAAGATTA 1087
DB 1033 ATGAGTACGGAACCACTTTCAATATCCGTGAAATTTGATTAAGATTA 1092
QY 1088 ATGATTAATCCACTACATGTCCTGACACAGTAAACCGTATTTGAGTCCAGAGATGAAC 1147
DB 1093 ATGATTAATCCACTACATGTCCTGACACAGTAAACCGTATTTGAGTCCAGAGATGAAC 1152
QY 1148 GACTGGGTAAACAGATGTGGGACCCCTTACCTGACATGACAGAGATGAAGAAATTTCTCCA 1207
DB 1153 GACTGGGTAAACAGATGTGGGACCCCTTACCTGACATGACAGAGATGAAGAAATTTCTCCA 1212
QY 1208 ACAAGTTTCTTAACTACAGAGATGTGGAGCAAACTCCCAAACTTTCCAGATGATGACTCT 1267
DB 1213 ACAAGTTTCTTAACTACAGAGATGTGGAGCAAACTCCCAAACTTTCCAGATGATGACTCT 1272
QY 1268 TCCTAATCCAAACCTATGTCTGGAATGTTTACAGTTAGCTTAAACAGTCTTGAAGAGCAG 1327
DB 1273 TCCTAATCCAAACCTATGTCTGGAATGTTTACAGTTAGCTTAAACAGTCTTGAAGAGCAG 1332
QY 1328 ATACTCTCTGATCAACTTAAACATATAGATGTCTGACAAAGATTTTCAACACCTTTGTT 1387
DB 1333 ATACTCTCTGATCAACTTAAACATATAGATGTCTGACAAAGATTTTCAACACCTTTGTT 1392
QY 1388 TTGTGCAAAATCAAGTATTTGATATCAATGATCAAGATCCCATCTTTGAAAAATCAGATT 1447
DB 1393 TTGTGCAAAATCAAGTATTTGATATCAATGATCAAGATCCCATCTTTGAAAAATCAGATT 1452
QY 1448 ATGGAACCTGACTCTTGTCTGAAGACAAACATTTGGGTCCACATCTTAAACATCCAGG 1507
DB 1453 ATGGAACCTGACTCTTGTCTGAAGACAAACATTTGGGTCCACATCTTAAACATCCAGG 1512
QY 1508 CCACTGATGTGATGAGGCAATTTTACTGGGAGTTCTTAAATTTCTGTATCATATCTAAAG 1567
DB 1513 CCACTGATGTGATGAGGCAATTTTACTGGGAGTTCTTAAATTTCTGTATCATATCTAAAG 1572
QY 1568 GAGACAGTGAAGGAGCGCTGGGGTTTGAACACAGATCCCATACCAACACCGGATATGTCA 1627
DB 1573 GAGACAGTGAAGGAGCGCTGGGGTTTGAACACAGATCCCATACCAACACCGGATATGTCA 1632
QY 1628 TAAATTAAGAGCTCTTGAATTTTGAACACAGACCTGTTTCCACATTTGTGTTCAAGCAG 1687
DB 1633 TAAATTAAGAGCTCTTGAATTTTGAACACAGACCTGTTTCCACATTTGTGTTCAAGCAG 1692
QY 1688 AAAATCTGAGCCTTGAATTTTGTGTGTGTAAGTCAATGCAATGTTCTTTTGCAGATTTCA 1747
DB 1693 AAAATCTGAGCCTTGAATTTTGTGTGTGTAAGTCAATGCAATGTTCTTTTGCAGATTTCA 1752
QY 1748 CGCTTATTTGACAGATGTGAATGAAGACCTCAATTTTCCAAACAGTATTCAGAGCA 1807
DB 1753 CGCTTATTTGACAGATGTGAATGAAGACCTCAATTTTCCAAACAGTATTCAGAGCA 1812
QY 1808 AAGTCAGTGAAGATGTATAGCACTTAAAGTGGGCAATGTGACTGCCAAGATTCAG 1867

DB 1813 AAGTCAGTGAAGATGTATAGCACTTAAAGTGGGCAATGTGACTGCCAAGATTCAG 1872
QY 1868 AAGTCTGACATTAAGCTATTTCACTGAGGGGACACCAAGGTGGCTTAAATTAAGC 1927
DB 1873 AAGTCTGACATTAAGCTATTTCACTGAGGGGACACCAAGGTGGCTTAAATTAAGC 1932
QY 1928 ACGTGAATGTGAGATCTTTAGTGTGCTCATTTGACAGAGAACCGGAAGTCCATATC 1987
DB 1933 ACGTGAATGTGAGATCTTTAGTGTGCTCATTTGACAGAGAACCGGAAGTCCATATC 1992
QY 1988 GGGTACAGATGTGGGCGCACAGAGTGAAGGGGGTCTTCTTAACTGTGTCAAGATTCC 2047
DB 1993 GGGTACAGATGTGGGCGCACAGAGTGAAGGGGGTCTTCTTAACTGTGTCAAGATTCC 2052
QY 2048 ACGTGAATCTTATGATGTGAATGAACCACTTCCAGAGTACGCCAAGATTAACGGGCT 2107
DB 2053 ACGTGAATCTTATGATGTGAATGAACCACTTCCAGAGTACGCCAAGATTAACGGGCT 2112
QY 2108 TGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTCAAGGCTAATGATGATG 2167
DB 2113 TGTCTTCTGCAATCCCTCAGTGCACCTGGAAGTCTCATTTTCAAGGCTAATGATGATG 2172
QY 2168 ATCAGCACTTATTTGGGGTCCCATTTTATCATTTTCCCTGGCAGTGAAGCTTACAA 2227
DB 2173 ATCAGCACTTATTTGGGGTCCCATTTTATCATTTTCCCTGGCAGTGAAGCTTACAA 2232
QY 2228 ACGACTGGGAAGCTTCCAAATTCATATGATATCCGCACTGTCTACAGGCAACAG 2287
DB 2233 ACGACTGGGAAGCTTCCAAATTCATATGATATCCGCACTGTCTACAGGCAACAG 2292
QY 2288 ACTTATGAGAGAGGGGATATGTCGTGATCCGATCAATGATGGGGGTGGCCACCT 2347
DB 2293 ACTTATGAGAGAGGGGATATGTCGTGATCCGATCAATGATGGGGGTGGCCACCT 2352
QY 2348 TGGAAAGCAATGTTTCTTTTACAGATTACATTTGCAAGTGTGTGAGAGAAAGTTGTTCC 2407
DB 2353 TGGAAAGCAATGTTTCTTTTACAGATTACATTTGCAAGTGTGTGAGAGAAAGTTGTTCC 2412
QY 2408 GGCACAGAGTACACGAGTGGGATACCACTGTGGGCAATGGAGTTGTATCTGCTGA 2467
DB 2413 GGCACAGAGTACACGAGTGGGATACCACTGTGGGCAATGGAGTTGTATCTGCTGA 2472
QY 2468 CCACCTCTGAGGATTTGGATTAATTTTGAAGTGTGATTAACCGATTAAGAGATTA 2527
DB 2473 CCACCTCTGAGGATTTGGATTAATTTTGAAGTGTGATTAACCGATTAAGAGATTA 2532
QY 2528 AAGGCAAGATTAATTTGAAGTGTCAAGCATCTGAAATCAAACTGTGAGAGCTGAA 2587
DB 2533 AAGGCAAGATTAATTTGAAGTGTCAAGCATCTGAAATCAAACTGTGAGAGCTGAA 2592
QY 2588 TTTGAAAGAGATTTGAAATTTATATGACAGTGTATTTGACACAAACATCTGATC 2647
DB 2593 TTTGAAAGAGATTTGAAATTTATATGACAGTGTATTTGACACAAACATCTGATC 2652
QY 2648 CTATTAATCTTCACTAAGGATTAATTTTAAACAGATTAATCTCTGTGCTT 2707
DB 2653 CTATTAATCTTCACTAAGGATTAATTTTAAACAGATTAATCTCTGTGCTT 2712
QY 2708 TAAATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTGTCTGTGCGCCAGGCTGAG 2767
DB 2713 TAAATTTGCTAAATATTTCTTTTGAAGTGAAGTCTTGTCTGTGCGCCAGGCTGAG 2772
QY 2768 TACAGTGTGTGATCCCACTACCTGCAACCTTCCGCTCTGTGGTTCAATGATTTCTCT 2827
DB 2773 TACAGTGTGTGATCCCACTACCTGCAACCTTCCGCTCTGTGGTTCAATGATTTCTCT 2832
QY 2828 GCTCAGCTTCCATAGTATGAGGTTTACAGGACCCACACAGTCCAGCTAATTTT 2887
DB 2833 GCTCAGCTTCCATAGTATGAGGTTTACAGGACCCACACAGTCCAGCTAATTTT 2892
QY 2888 GTATTTTAAAGAGCGGGTTTGCATTTTGGCAGGCTGTGTAATCTTCAAGCT 2947

Db 2893 GTATTTTATAGAGAGGGGTTTCGCATTGGCCAGGCTGGCTTGAACCTCTAGCGT 2952
Qy 2948 CAAAGTATCTGCTGCTGGTGGTCTCCCAATACAGGCTGAACCACTGCACTCACTT 3007
Db 2953 CAAAGTATCTGCTGCTGGTGGTCTCCCAATACAGGCTGAACCACTGCACTCACTT 3012
Qy 3008 AGATATTTCAATGCTATACACATTAAGAGATTTTTCATTTTCCATGACATTTTCC 3067
Db 3013 AGATATTTCAATGCTATACACATTAAGAGATTTTTCATTTTCCATGACATTTTCC 3072
Qy 3068 CTGCGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3127
Db 3073 CTGCGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3132
Qy 3128 TATTCGTACATTTTCTTTTATCAGAGATATATCAGTGTGCTCATAGACCTGCT 3187
Db 3133 TATTCGTACATTTTCTTTTATCAGAGATATATCAGTGTGCTCATAGACCTGCT 3192
Qy 3188 GGAATTCATTTATGTTTTTCTGATTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 3247
Db 3193 GGAATTCATTTATGTTTTTCTGATTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCT 3252
Qy 3248 GTATTTCTAGATTTTCAACATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3307
Db 3253 GTATTTCTAGATTTTCAACATTTGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
Qy 3308 AAAATAAATAAG 3333
Db 3313 AAAATAAATAAG 3338

RESULT 7
US-09-880-107-3807
; Sequence 3807, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scheff, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 3807
; LENGTH: 3654
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X83228
US-09-880-107-3807

Query Match 99.1%; Score 3314.8; DB 10; Length 3654;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3319; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 188 TGACATTTTCTATTTATGAGAGGCAAGAACCGAGTCAATATATTCAGTTTAAAGGCA 247
Db 193 TGACATTTTCTATTTATGAGAGGCAAGAACCGAGTCAATATATTCAGTTTAAAGGCA 252
Qy 248 ATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 307
Db 253 ATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
Qy 308 GGGAGGAGCTTCTGATTTTCAACAGAGCTTTGAGACGGGAGAACATATCTACCAATC 367
Db 313 GGGAGGAGCTTCTGATTTTCAACAGAGCTTTGAGACGGGAGAACATATCTACCAATC 372
Qy 368 TCCAGGTTGAGGCTGAGAGGCTAATGAGATTTAGTGAAGGCTCCATGACCA 427
Db 373 TCCAGGTTGAGGCTGAGAGGCTAATGAGATTTAGTGAAGGCTCCATGACCA 432
Qy 428 TAGAGTGAAGGACATCAACGATCGAACCGATTTCTCAGTCAAGATGAGAGGCT 487
Db 433 TAAAGTGAAGGACATCAACGATCGAACCGATTTCTCAGTCAAGATGAGAGGCT 492
Qy 488 CAGTAAGGAGAACTCTGCGCCAGAGAAAGCCCTTCTGATGATGATGATGATGATGATGAT 547
Db 493 CAGTAAGGAGAACTCTGCGCCAGAGAAAGCCCTTCTGATGATGATGATGATGATGATGAT 552
Qy 548 ATGATCCGGGCACTCCCAATGGCCAGCTTTATTAACAGATTTGATGATGATGATGATGAT 607
Db 553 ATGATCCGGGCACTCCCAATGGCCAGCTTTATTAACAGATTTGATGATGATGATGATGAT 612
Qy 608 TCAACATGTCATGATCTTCAAGATCAACCAACAAACGGAGGCACTCTCTTACCGGAG 667
Db 613 TCAACATGTCATGATCTTCAAGATCAACCAACAAACGGAGGCACTCTCTTACCGGAG 672
Qy 668 AGGATCTCAGAAATTTGAATCTGCTTAAGAAATCTTCTTATATCTGATGATGATGATGAT 727
Db 673 AGGATCTCAGAAATTTGAATCTGCTTAAGAAATCTTCTTATATCTGATGATGATGATGAT 732
Qy 728 AGGATCGGGAGGCGAGAGTGAATCTCTCAGTATACCAATCTGATGATGATGATGATGAT 787
Db 733 AGGATCGGGAGGCGAGAGTGAATCTCTCAGTATACCAATCTGATGATGATGATGATGAT 792
Qy 788 TGACAGAGATATTTGGAGAGACCAACAAACCTGTGAGATGATGAGAACTCAACTGATC 847
Db 793 TGACAGAGATATTTGGAGAGACCAACAAACCTGTGAGATGATGAGAACTCAACTGATC 852
Qy 848 CTGACCCCATCAAAATCCTCAGAGTGGGTGAGATATCCGGTGCACATATTCCTTAG 907
Db 853 CTGACCCCATCAAAATCCTCAGAGTGGGTGAGATATCCGGTGCACATATTCCTTAG 912
Qy 908 TTGACAAAGAGAGCTGCGCAAGATTCCTTCAATTTCAATGACAGAGAGAGATATTTACG 967
Db 913 TTGACAAAGAGAGCTGCGCAAGATTCCTTCAATTTCAATGACAGAGAGAGATATTTACG 972
Qy 968 TGACTGAGCCCTTGAACCGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 1027
Db 973 TGACTGAGCCCTTGAACCGAGAGAGAGAGAGATGATGATGATGATGATGATGATGATGAT 1032
Qy 1028 ATGAGTACGAG 1087
Db 1033 ATGAGTACGAG 1092
Qy 1088 ATGATATTCAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1147
Db 1093 ATGATATTCAGACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
Qy 1148 GACTGGTAAACAGTATGAG 1207
Db 1153 GACTGGTAAACAGTATGAG 1212
Qy 1208 ACAATTTCTTAACTACAGATTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1267
Db 1213 ACAATTTCTTAACTACAGATTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1272

QY	1266	TCCTATCCAAACCTATGCTGGAAAGTTTACAGTTAGCTTAAAGCTCTTGAAGAACAG	132
Db	1273	TCCTATCCAAACCTATGCTGGAAAGTTTACAGTTAGCTTAAAGCTCTTGAAGAACAG	133
QY	1328	ATACTCTTCAGTACACTTAAAGATAGAGGTCTGACAAAGATTTCAGAACCTTTGTT	1387
Db	1333	ATACTCTTCAGTACACTTAAAGATAGAGGTCTGACAAAGATTTCAGAACCTTTGTT	1392
QY	1388	TTGTGCAAAATCAACGTTATTGATATCATGATGATGCCATCTTTGAAAAATCAGATT	1447
Db	1393	TTGTGCAAAATCAACGTTATTGATATCAAGATGATGCCATCTTTGAAAAATCAGATT	1452
QY	1448	ATGGAACCTGACTCTGTGCTGAAGACAAATATGGGGTCCACATCTTAAACATCCAG	1507
Db	1453	ATGGAACCTGACTCTGTGCTGAAGACAAATATGGGGTCCACATCTTAAACATCCAG	1512
QY	1508	CCACTGATGCTGATGAGCCATTACTGGAGTCTTAAATTTGTATCATATCATAAAG	1566
Db	1513	CCACTGATGCTGATGAGCCATTACTGGAGTCTTAAATTTGTATCATATCATAAAG	1572
QY	1568	GAGACAGTGAAGGACGCTGGGGGTTGACACAGATCCCCATACAAACCGGATATGTCA	1627
Db	1573	GAGACAGTGAAGGACGCTGGGGGTTGACACAGATCCCCATACAAACCGGATATGTCA	1632
QY	1628	TAAATTAAGAAAGCCTCTTGATTTTGAACACAGCGCTGTTCCAAATGTGTTCAAAGC	1687
Db	1633	TAAATTAAGAAAGCCTCTCTGATTTTGAACACAGCGCTGTTCCAAATGTGTTCAAAGC	1692
QY	1688	AAATCTGAGGCTCTAGTGTTTGGTGTGTAAGTCAATGCAAGTCTTTGGCAAGTCA	1747
Db	1693	AAATCTGAGGCTCTAGTGTTTGGTGTGTAAGTCAATGCAAGTCTTTGGCAAGTCA	1752
QY	1748	CGCTTATTGTGACAGATGTGAATGAAGACCTCAATTTTCCCAACAGTATTCGAAGCA	1807
Db	1753	CGCTTATTGTGACAGATGTGAATGAAGACCTCAATTTTCCCAACAGTATTCGAAGCA	1812
QY	1808	AAGTCAGTGAAGATTTAGCTATAGGCACTTAAAGTGGCGATGTACTGCCAAGATCCAG	1866
Db	1813	AAGTCAGTGAAGATTTAGCTATAGGCACTTAAAGTGGCGATGTACTGCCAAGATCCAG	1872
QY	1868	AAGTCTGGAACATAAGCTATTACCTGAGGGGAGACACAAGAGGTTGGCTTAAATTTGACC	1927
Db	1873	AAGTCTGGAACATAAGCTATTACCTGAGGGGAGACACAAGAGGTTGGCTTAAATTTGACC	1932
QY	1928	ACGTGACTGTGAGATCTTTAGTGTGCTCAATTGACAGAGAACCCGGAAGTCCATATC	1987
Db	1933	ACGTGACTGTGAGATCTTTAGTGTGCTCAATTGACAGAGAACCCGGAAGTCCATATC	1992
QY	1988	GGGTACAGTGTGTGCCACACAGATGAGGGGGTCTTCTTAAAGCTCTGTGTGACAGTTCC	2047
Db	1993	GGGTACAGTGTGTGCCACACAGATGAGGGGGTCTTCTTAAAGCTCTGTGTGACAGTTCC	2052
QY	2048	ACCTGATCTCTTATGAGATGGAATGACAAACCTCCGAGGCTGACGAAAGACTACAGGGCT	2107
Db	2053	ACCTGATCTCTTATGAGATGGAATGACAAACCTCCGAGGCTGACGAAAGACTACAGGGCT	2112
QY	2108	TGTTCTTGTGCATCCCTCAGTGCACCTGGAAGTCTGATTTCCAGGCTTACTGATGATG	2167
Db	2113	TGTTCTTGTGCATCCCTCAGTGCACCTGGAAGTCTGATTTCCAGGCTTACTGATGATG	2172
QY	2168	ATCAGCACTTATTTGGGGTCCCATTTTTCATTTTCCCTGGGCAGTGAAAGCTTACAA	2227
Db	2173	ATCAGCACTTATTTGGGGTCCCATTTTTCATTTTCCCTGGGCAGTGAAAGCTTACAA	2232
QY	2228	ACGACTGGGAAGTTTCCAAAAATCAATGTGATCTCAGGCGGACTGTCTACAGGCAACAG	2287
Db	2233	ACGACTGGGAAGTTTCCAAAAATCAATGTGATCTCAGGCGGACTGTCTACAGGCAACAG	2292
QY	2288	ACTTTGAGAGAGGGCGTATGTGCTCTTGTATCCGATCCGATCAATGATGGGGGTGGCCACCT	2347
Db	2293	AGTTTGAAGAGGGGATGATGTGCTGCTTGTATCCGATCAATGATGGGGGTGGCCACCT	2352
QY	2348	TGGAAGCATGTGTTCTTTACACAGTTATCTGCACTTGTGTGGAAGAAAGTTGTTTCC	2407

Db	2353	TGANGGATTTGTTCTTTTACCACTTACATTCGTGAGTTGTGAGAAAGATTGTTCC	2412
Qy	2408	GGCCAGCAGGTACACAGACTGGAGTACCACCTGTGGCAGTGGCAGTTGGTACTGCTGA	2467
Db	2413	GGCCAGCAGGTACACAGACTGGAGTACCACCTGTGGGAGTGGCAGTTGGTACTGCTGA	2472
Qy	2468	CCACCCCTCTGCGTATTTGGTATAATTTTGAAGATTTGGTTTATTCGCTATAAAGAAAGTA	2522
Db	2473	CAACCCCTCTGCGTATTTGGTATAATTTTGAAGATTTGGTTTATTCGCTATAAAGAAAGTA	2532
Qy	2528	AAGCAAAAGATATGTTGAAAAGTGCCTCAGCATGTGAAGTCAAACTCTGAGAACTGAA	2587
Db	2533	AAGCAAAAGATATGTTGAAAAGTGCCTCAGCATGTGAAGTCAAACTCTGAGAACTGAA	2592
Qy	2588	TTTGAAGAAAGATGTTGAAATTTATATATGCAAGTCTATTTCAGACACACCATCTCATC	2647
Db	2593	TTTGAAGAAAGATGTTGAAATTTATATATGCAAGTCTATTTCAGACACACCATCTCATC	2655
Qy	2648	CTATTACTTTTCATCTCAACGTCATATTAATTTTAAACAGATATTCCTCTGCTCTT	2707
Db	2653	CTATTACTTTTCATCTCAACGTCATATTAATTTTAAACAGATATTCCTCTGCTCTT	2712
Qy	2708	TAAATATTTGCTAAATATTTCTTTTTTGAAGTGAAGTCTGTCTGTGTGGCCAGGCTGGAG	2767
Db	2713	TAAATATTTGCTAAATATTTCTTTTTTGAAGTGAAGTCTGTCTGTGTGGCCAGGCTGGAG	2772
Qy	2768	TACAGTGGTGTGATCCAGCTCATCGAACCTCGGCTCTCGGGTTACATGATTTCTCT	2827
Db	2773	TACAGTGGTGTGATCCAGCTCATCGAACCTCGGCTCTCGGGTTACATGATTTCTCT	2832
Qy	2828	GCTCAGCTTCTTAAGTAGCTGGGTTTACAGGCAACCAACCAACATGCCCAGCTAATTTTT	2887
Db	2833	GCTCAGCTTCTTAAGTAGCTGGGTTTACAGGCAACCAACCAACATGCCCAGCTAATTTTT	2892
Qy	2888	GTAATTTTTTAATAGAGAGGGGGTTTCGCTATTTGGCCAGGCTGGTCTTGAATCTCTGAAGT	2947
Db	2893	GTAATTTTTTAATAGAGAGGGGGTTTCGCTATTTGGCCAGGCTGGTCTTGAATCTCTGAAGT	2952
Qy	2948	CAAGTGAATCTGCTGCGCTTGCTGCCCAATACAGAGCAAGAAACAACCTGACCTACTT	3007
Db	2953	CAAGTGAATCTGCGCTTGCTGCCCAATACAGAGCAATGAACCACTGACCACTACTT	3012
Qy	3008	AGATATTTTCATGTGCTATAGACATTAAGAGATTTTTCATTTTCCATGACATTTTTCT	3067
Db	3013	AGATATTTTCATGTGCTATAGACATTAAGAGATTTTTCATTTTCCATGACATTTTTCT	3072
Qy	3068	CTCTGCAAAAGGCTAGCTACTGTGTGTTTTCCCTTTGGGGCAAGACAGCTATTAATAA	3127
Db	3073	CTCTGCAAAAGGCTAGCTACTGTGTGTGTTTTCCCTTTGGGGCAAGACAGCTATTAATAA	3132
Qy	3128	TATTCCTGACATTTTTCTTATACAGAGATATATCAGTGTGTCTCATGAACCTGCT	3187
Db	3133	TATTCCTGACATTTTTCTTATACAGAGATATATCAGTGTGTCTCATGAACCTGCT	3192
Qy	3188	GGATTCGATTAAGTTTTTTTCTGATTCGATTCCTGTGCCCTTATCTCTGACTCTCTTG	3247
Db	3193	GGATTCGATTAAGTTTTTTTCTGATTCATTCCTGTGTGCCCTTATCTCTGACTCTCTTG	3252
Qy	3248	GTATTTCACAGAAATTTCAAAATTTGTTCAGAGAAAGAAAGTGAAGACTCAGAGAAAT	3307
Db	3253	GTATTTCACAGAAATTTCAAAATTTGTTCAGAGAAAGAAAGTGAAGACTCAGAGAAAT	3312
Qy	3308	AAATTAATTAAGAACGCTTTTTCG 3333	
Db	3313	AAATTAATTAAGAACGCTTTTTCG 3338	

RESULT 8
US-10-025-380-1086
; Sequence 1086, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:


```

|||||
Db 2125 ACTAAATGGGCATGTGACGCGAAGATCCAAAGGTCTGGACATAGCTATTCACTG 2184
|||
Qy 1894 AGGGGAGACACAAGAGGTTGGCTTAAATGACACGTCGTGGTATCTTTAGTGTG 1953
|||
Db 2185 AGGGGAGACACAAGAGGTTGGCTTAAATGACACGTCGTGGTATCTTTAGTGTG 2244
|||
Qy 1954 GCTTCATTGACACAGAGAGCGGAGTCAATATCGGGTACAGTGTGGCACAGAGTA 2013
|||
Db 2245 GCTTCATTGACACAGAGAGCGGAGTCAATATCGGGTACAGTGTGGCACAGAGTA 2304
|||
Qy 2014 GGGGGGCTTCTTAACTCTGTGTCAAGTTCACCTGATCCTTAGATGTGATGAC 2073
|||
Db 2305 GGGGGGCTTCTTAACTCTGTGTCAAGTTCACCTGATCCTTAGATGTGATGAC 2364
|||
Qy 2074 AACCTCCAGAGGTCAGACAGAGTACACGGGCTGTCTTCCGCAATCCCTCAGTGA 2133
|||
Db 2365 AACCTCCAGAGGTCAGACAGAGTACACGGGCTGTCTTCCGCAATCCCTCAGTGA 2424
|||
Qy 2134 CCGAAGTCTCAATTTGAGAGGCTAGATGATGATGACACTAATTCGGGGTCCCAT 2193
|||
Db 2425 CCGAAGTCTCAATTTGAGAGGCTAGATGATGATGACACTAATTCGGGGTCCCAT 2484
|||
Qy 2194 TTTACATTTTCCCTCGGAGTGAAGCTTACAAAACGACTGGGAAATTTCCAAATCAAT 2253
|||
Db 2485 TTTACATTTTCCCTCGGAGTGAAGCTTACAAAACGACTGGGAAATTTCCAAATCAAT 2544
|||
Qy 2254 GGTACTCATGCGCGAGTGTCTACACGAGCAGACACTTGAAGAGAGGGGTATGTCGTC 2313
|||
Db 2545 GGTACTCATGCGCGAGTGTCTACACGAGCAGACACTTGAAGAGAGGGGTATGTCGTC 2604
|||
Qy 2314 TTGATCGCATCAATGATGAGGGGTGCGCACCCCTGGAGAGCAATGTTCTTTACAGT 2373
|||
Db 2605 TTGATCGCATCAATGATGAGGGGTGCGCACCCCTGGAGAGCAATGTTCTTTACAGT 2664
|||
Qy 2374 ACATTTCTGCACTGTGTGTGAAGAGTGTGTTCCGCGCAGCAGCTCACAGACTGGAGTA 2433
|||
Db 2665 ACATTTCTGCACTGTGTGTGAAGAGTGTGTTCCGCGCAGCAGCTCACAGACTGGAGTA 2724
|||
Qy 2434 CCACAGTGGGCGAAGGAGTGTGATCTGCTGACACCCCTTCTGGATGTGTATATAT 2493
|||
Db 2725 CCACAGTGGGCGAAGGAGTGTGATCTGCTGACACCCCTTCTGGATGTGTATATAT 2784
|||
Qy 2494 TTAGCAGTGTGTGTTATCCGCATTAAGAGATTAAGGCAAGTATATGTTGAAGTCT 2553
|||
Db 2785 TTAGCAGTGTGTGTTATCCGCATTAAGAGATTAAGGCAAGTATATGTTGAAGTCT 2844
|||
Qy 2554 CAAGCATCTGAAGTCAAACTCTGAGAGCTGA 2586
|||
Db 2845 CAAGCATCTGAAGTCAAACTCTGAGAGCTGA 2877

```

RESULT 9

```

US-09-922-217-1086
; Sequence 1086, Application US/09922217
; Patent No. US2002007641A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yudi
; APPLICANT: Smith, Carole Lynn
; APPLICANT: Wang, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217

```

```

; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1086
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-1086

Query Match      72.7%; Score 2432.6; DB 10; Length 2877;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 154 CAAGAGGGGAAAGTTAGTGACCCCTGAACCCATGACATTTCTATTATGAAGGCCAA 213
Db 445 CAAAGAGGGGAAAGTTAGTGACCCCTGAACCCATGACATTTCTATTATGAAGGCCAA 504
Qy 214 GAAACGAGTCAATTAATTCAGTTTAAGGCCAATCTCTGCTGTGACTTTGAAGTA 273
Db 505 GAAACGAGTCAATTAATTCAGTTTAAGGCCAATCTCTGCTGTGACTTTGAAGTA 564
Qy 274 ACTGGGAGACAGACAAATTTTGTATGAACGGGAGGACTTCTATTACAAACAGA 333
Db 565 ACTGGGAGACAGACAAATTTTGTATGAACGGGAGGACTTCTATTACAAACAGA 624
Qy 334 GCCTTGAACAGGAAACAAGATCTACTCAATCTCAGGTTGACGCCCTGAGCGTAAAT 393
Db 625 GCCTTGAACAGGAAACAAGATCTACTCAATCTCAGGTTGACGCCCTGAGCGTAAAT 684
Qy 394 GGAATTAATAGTGAGGAGTCCAGTCCCTATCACCATAAGAGTGAAGACATCAACAT 453
Db 685 GGAATTAATAGTGAGGAGTCCAGTCCCTATCACCATAAGAGTGAAGACATCAACAT 744
Qy 454 CGAACCAAGTTTCTCAGTCAAAAGTACGAAGCTCAATGAAGCAGAACTCTGCCAGGA 513
Db 745 CGAACCAAGTTTCTCAGTCAAAAGTACGAAGCTCAATGAAGCAGAACTCTGCCAGGA 804
Qy 514 AAGCCCTCTGTATGATGCAATGCGACAGACCTGAGATCCGGGCCCTCCCAATGGCCAG 573
Db 805 AAGCCCTCTGTATGATGCAATGCGACAGACCTGAGATCCGGGCCCTCCCAATGGCCAG 864
Qy 574 CTTTATTAACAGATTTGATCCAGCTTCCATGATCAACATGTACTACTTACAGATC 633
Db 865 CTTTATTAACAGATTTGATCCAGCTTCCATGATCAACATGTACTACTTACAGATC 924
Qy 634 AACCAAAAACGGAGCCATCTCTTTACCCGAGAGGATCTCAGAAATTTGAATCTGCT 693
Db 925 AACCAAAAACGGAGCCATCTCTTTACCCGAGAGGATCTCAGAAATTTGAATCTGCT 984
Qy 694 AAGAAATCTCTCTATTAATGTGATCTCAGTGAAGAGACATGGGAGCCAGAGTGAAT 753
Db 985 AAGAAATCTCTCTATTAATGTGATCTCAGTGAAGAGACATGGGAGCCAGAGTGAAT 1044
Qy 754 TCTTCAGTATACCATCTGTGATATCATATGACAGAGAAATTTGAAAGCACCA 813
Db 1045 TCTTCAGTATACCATCTGTGATATCATATGACAGAGAAATTTGAAAGCACCA 1104
Qy 814 AAAACCTGTGAGATGGTGAAGAACTCACTGATCTCACCCTCAATCAATCACTCAGTGTG 873
Db 1105 AAAACCTGTGAGATGGTGAAGAACTCACTGATCTCACCCTCAATCAATCACTCAGTGTG 1164
Qy 874 CGGTGAATGATCCGGGTGACATATTTCTGATGTTGCAAGAGAGTGGCAAGATTC 933
Db 1165 CGGTGAATGATCCGGGTGACATATTTCTGATGTTGCAAGAGAGTGGCAAGATTC 1224
Qy 934 CCATTTCAATTAAGACAGAGAGATTTTACGTGACTCAGCCCTTGGACCGAGAGAA 993
Db 1225 CCATTTCAATTAAGACAGAGAGATTTTACGTGACTCAGCCCTTGGACCGAGAGAA 1284
Qy 994 AAGAGTCAATGTTTATGACAGTTGCAAGATGATGATGAGAAACCACTTTCATAT 1053
Db 1285 AAGAGTCAATGTTTATGACAGTTGCAAGATGATGATGAGAAACCACTTTCATAT 1344

```

QY 1054 CCGCTGGAATTCATGTAATAAGTAAATTAATGATATACCACTACATGTCGTC 1113
 DB 1345 CCGCTGGAATTCATGTAATAAGTAAATTAATGATATACCACTACATGTCGTC 1404
 QY 1114 CCAAGTAACCGTATTTAGAGTCCAGAGATGAAAGATGGGTAAAGTATCGGAGCCCTT 1173
 DB 1405 CCAAGTAACCGTATTTAGAGTCCAGAGATGAAAGATGGGTAAAGTATCGGAGCCCTT 1464
 QY 1174 ACTGCATATGACAGAGGATGAAAGAAATATCTGCCACAGTTTCTTAAATCTACAGATTTGTG 1233
 DB 1465 ACTGCATATGACAGAGGATGAAAGAAATATCTGCCACAGTTTCTTAAATCTACAGATTTGTG 1524
 QY 1234 GACCAAACTCCCAACTCCCATGATGATGATCTCTCTTAATCCAAACTCTATGCGGAATG 1293
 DB 1525 GACCAAACTCCCAACTCCCATGATGATGATCTCTCTTAATCCAAACTCTATGCGGAATG 1584
 QY 1294 TTACAGTTAGCTAAACAGTCTCTTGAAGACAGATATCTCTCTGATCAACCTTAACGATA 1353
 DB 1585 TTACAGTTAGCTAAACAGTCTCTTGAAGACAGATATCTCTCTGATCAACCTTAACGATA 1644
 QY 1354 GAGGTGTCTGACAAAGATTTCAAGACCCCTTTGTTGTGCAATCAACGTTATTTGATATC 1413
 DB 1645 GAGGTGTCTGACAAAGATTTCAAGACCCCTTTGTTGTGCAATCAACGTTATTTGATATC 1704
 QY 1414 AATGATCAGATCCCATCTTTGAAATTCAGATTAATGAAACCTGATCTTGCGTAAGAC 1473
 DB 1705 AATGATCAGATCCCATCTTTGAAATTCAGATTAATGAAACCTGATCTTGCGTAAGAC 1764
 QY 1474 ACAAACTTTGGGTCCACATCTTTAACATCCAGGCGCATGATGATGAGCCATTTACT 1533
 DB 1765 ACAAACTTTGGGTCCACATCTTTAACATCCAGGCGCATGATGATGAGCCATTTACT 1824
 QY 1534 GGGAGTTCTAAATTTCTGTATCATATCAATTAAGGGAAGACAGTGAAGGAGCCCTGGGGGTT 1593
 DB 1825 GGGAGTTCTAAATTTCTGTATCATATCAATTAAGGGAAGACAGTGAAGGAGCCCTGGGGGTT 1884
 QY 1594 GACACATATCCCATCCATCCACACCGGATATGTCATTAATTAAGGCGCTTGATTTTGA 1653
 DB 1885 GACACATATCCCATCCATCCACACCGGATATGTCATTAATTAAGGCGCTTGATTTTGA 1944
 QY 1654 ACAGCAGCTGTTTCCACATGTTGTTCAAGACAGAAATCTGAGCCCTGATGTTTGT 1713
 DB 1945 ACAGCAGCTGTTTCCACATGTTGTTCAAGACAGAAATCTGAGCCCTGATGTTTGT 2004
 QY 1714 GTGAATGATCAATGCAAGTTCTTTTCCAGTTTCCAGTTTATGTCAGATGTAATGAA 1773
 DB 2005 GTGAATGATCAATGCAAGTTCTTTTCCAGTTTCCAGTTTATGTCAGATGTAATGAA 2064
 QY 1774 GCACCTCAATTTTCCCAACAGTATTCACAGGAAAGTCAAGTGAATGATAGTATGAGC 1833
 DB 2065 GCACCTCAATTTTCCCAACAGTATTCACAGGAAAGTCAAGTGAATGATAGTATGAGC 2124
 QY 1834 ACTAAATGGGCAATGTCAGTCCAGAGATCCAGAAAGTCTGACATTAAGCTATTTCACTG 1893
 DB 2125 ACTAAATGGGCAATGTCAGTCCAGAGATCCAGAAAGTCTGACATTAAGCTATTTCACTG 2184
 QY 1894 AGGGGAGACACAAAGGTTGGCTTAAATTAATGACACGTCAGTCTGATGATTTTATGTTG 1953
 DB 2185 AGGGGAGACACAAAGGTTGGCTTAAATTAATGACACGTCAGTCTGATGATTTTATGTTG 2244
 QY 1954 GCTTCATTTGACAGAGAAAGCCGGAAGTCCATATGCGGATCAAGAGTGGCCACAGAAATA 2013
 DB 2245 GCTTCATTTGACAGAGAAAGCCGGAAGTCCATATGCGGATCAAGAGTGGCCACAGAAATA 2304
 QY 2014 GGGGGGCTTCTTAAAGTCTGTGTCAAGATTTCAACCTGATCTTATGATGATGATGAC 2073
 DB 2305 GGGGGGCTTCTTAAAGTCTGTGTCAAGATTTCAACCTGATCTTATGATGATGATGAC 2364
 QY 2074 AACCTTCCAGAGCTAGCCAGAGACTACACGGGCTTTGTTCTTCTGCAATCCCTCAAGTGA 2133
 DB 2365 AACCTTCCAGAGCTAGCCAGAGACTACACGGGCTTTGTTCTTCTGCAATCCCTCAAGTGA 2424
 QY 2134 CCTGGAAGTCTCATTTTCTGAGGCTATGATGATGATGACACTTAATTTCCGGGGTCCCAT 2193

DB 2425 CCGGAGAGCTCATTTTGGAGCTATGATGATACAGACATTAATTTGGGGTCCCAT 2484
 QY 2194 TTACATTTTCCCTCGGCGATGGAAGCTTAACAAAAGACTGGGAATTTCCAAATCAT 2253
 DB 2485 TTACATTTTCCCTCGGCGATGGAAGCTTAACAAAAGACTGGGAATTTCCAAATCAT 2544
 QY 2254 GGTACTCATGCCGACTGTCTACAGGACACAGACTTTGAGAGAGGCGTATGTCTC 2313
 DB 2545 GGTACTCATGCCGACTGTCTACAGGACACAGACTTTGAGAGAGGCGTATGTCTC 2604
 QY 2314 TTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAAGGATTTTCTTTACAGTT 2373
 DB 2605 TTGATCCGATCAATGATGAGGGGTGCGCCACCTTGAAGGATTTTCTTTACAGTT 2664
 QY 2374 ACATTTCTGAGTTGTGTGGAAGAAATGTTTCCGGCCAGCAGTCACAGACTGGGATA 2433
 DB 2665 ACATTTCTGAGTTGTGTGGAAGAAATGTTTCCGGCCAGCAGTCACAGACTGGGATA 2724
 QY 2434 CCCACTGTGGGCGATGGCAGTTGATATCTGACACACCTTCTGATGATGATTAAT 2493
 DB 2725 CCCACTGTGGGCGATGGCAGTTGATATCTGACACACCTTCTGATGATGATTAAT 2784
 QY 2494 TTACAGTTGTGTTATTCGCAATPAAGAGATPAAGGCAAGATATGTTGAAGTCT 2553
 DB 2785 TTACAGTTGTGTTATTCGCAATPAAGAGATPAAGGCAAGATATGTTGAAGTCT 2844
 QY 2554 CAAGCATCTGAAGTCAAACTCTGAGAGACTGA 2586
 DB 2845 CAAGCATCTGAAGTCAAACTCTGAGAGACTGA 2877

RESULT 10
 US-09-833-263-1086
 ; Sequence 1086, Application US/09833263
 ; Patent No. US20020110547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1086
 ; LENGTH: 2877
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(2877)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-263-1086
 Query Match 72.7%; Score 2432.6; DB 10; Length 2877;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2432; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 154 CAAGAGGGGAAGTTAGTGAACCCCGAAGAACCATGACATTTCTTATGAAAGGCAA 213
 DB 445 CAAGAGGGGAAGTTAGTGAACCCCGAAGAACCATGACATTTCTTATGAAAGGCAA 504
 QY 214 GAACGAGTCAATTAATTCAGTTTAAAGGCAATCTCTGCTGATGATTTGAAGTGA 273
 DB 505 GAACGAGTCAATTAATTCAGTTTAAAGGCAATCTCTGCTGATGATTTGAAGTGA 564
 QY 274 ACTGGGAGACAGACAAATATTTGTATGAAGCGGAGGAGACTTGTATTAACAAGCA 333
 DB 565 ACTGGGAGACAGACAAATATTTGTATGAAGCGGAGGAGACTTGTATTAACAAGCA 624

QY 334 GCGTTGACAGGAAACAAGATCTACTCAATCTCCAGGTGGAGCCCTGGACGCTAAT 393
DB 625 GCCTTGACAGGAAACAAGATCTACTCAATCTCCAGGTGGAGCCCTGGACGCTAAT 684
QY 394 GGAATTTAGTGAAGGGTCCAGTCCCTATCACTAAGAGTAAAGACATCAAGCAAT 453
DB 685 GGAATTTAGTGAAGGGTCCAGTCCCTATCACTAAGAGTAAAGACATCAAGCAAT 744
QY 454 GCAACCAAGTTTCCAGTCAAGAGTCAAGAGGCTCAATTAAGGCAATCTCCGCCGGA 513
DB 745 GCAACCAAGTTTCCAGTCAAGAGTCAAGAGGCTCAATTAAGGCAATCTCCGCCGGA 804
QY 514 AAGCCCTTCTTGAATGTCAATGGCAGACCTGGATATCCGGCACTCCCAATGGCAG 573
DB 805 AAGCCCTTCTTGAATGTCAATGGCAGACCTGGATATCCGGCACTCCCAATGGCAG 864
QY 574 CTTTATTAACAGATTTGATTCAGTCCAGTCCCATGATCAATGTCTGATCTTCAATC 633
DB 865 CTTTATTAACAGATTTGATTCAGTCCAGTCCCATGATCAATGTCTGATCTTCAATC 924
QY 634 AACACAAACGGAGGCAATCTCTCTTAACCGAGAGGATCTCAAGAAATTGAATCCGCT 693
DB 925 AACACAAACGGAGGCAATCTCTCTTAACCGAGAGGATCTCAAGAAATTGAATCCGCT 984
QY 694 AAGAAATCTTCTAATATCTGATATCTCAAGTGAAGACATGGAGGCGAGAGTGAAT 753
DB 985 AAGAAATCTTCTAATATCTGATATCTCAAGTGAAGACATGGAGGCGAGAGTGAAT 1044
QY 754 TCCTTCAGTGAATCAACATCTGTGATATCAATGATCAAGAGATTTTGGAAAGCA 813
DB 1045 TCCTTCAGTGAATCAACATCTGTGATATCAATGATCAAGAGATTTTGGAAAGCA 1104
QY 814 AAACCTGTGAGATGTGGAAGAACTCAACGATCTTCAACCCCATCAAAATCACTCAGGTG 873
DB 1105 AAACCTGTGAGATGTGGAAGAACTCAACGATCTTCAACCCCATCAAAATCACTCAGGTG 1164
QY 874 CGGTGGAATATCCCGGTGCAATATCTTCTAGTTGACAAAGAGAGTCCCAAGATTC 933
DB 1165 CGGTGGAATATCCCGGTGCAATATCTTCTAGTTGACAAAGAGAGTCCCAAGATTC 1224
QY 934 CCAATTTCAATGACAGAGAGAGATATTTACGTGACTCAGCCCTTGAACCGAGAGAA 993
DB 1225 CCAATTTCAATGACAGAGAGAGATATTTACGTGACTCAGCCCTTGAACCGAGAGAA 1284
QY 994 AAGGATGCAATATTTTATGAGTTGCAAGAGTGAAGGAAACCACTTCAAT 1053
DB 1285 AAGGATGCAATATTTTATGAGTTGCAAGAGTGAAGGAAACCACTTCAAT 1344
QY 1054 CCGCTGGAATTTGATGTAAGTAAAGTAAATTAATGATTAATCACTACATGTCGCTCA 1113
DB 1345 CCGCTGGAATTTGATGTAAGTAAAGTAAATTAATGATTAATCACTACATGTCGCTCA 1404
QY 1114 CCAATTAACGTATTTGAGTCCAGAGATGAACGACTGGGTAAAGTATCGGACCTT 1173
DB 1405 CCAATTAACGTATTTGAGTCCAGAGATGAACGACTGGGTAAAGTATCGGACCTT 1464
QY 1174 ACTGCAATGACAGGATGAAGAAATATCTGCCAAGTTTCTTAACTACAGAGTTGTG 1233
DB 1465 ACTGCAATGACAGGATGAAGAAATATCTGCCAAGTTTCTTAACTACAGAGTTGTG 1524
QY 1234 GAGCAACTCCCAACTTCCATGATGAGCTCTTCTAATTCAAACCTATGCTGGAAG 1293
DB 1525 GAGCAACTCCCAACTTCCATGATGAGCTCTTCTAATTCAAACCTATGCTGGAAG 1584
QY 1294 TTAACATTAAGTAAAGAGTCTTGAAGAGCAAGTACTCTCAGTACCACTTAAGGATA 1353
DB 1585 TTAACATTAAGTAAAGAGTCTTGAAGAGCAAGTACTCTCAGTACCACTTAAGGATA 1644
QY 1354 GAGGTGTGACAAAGATTTCAAGACCTTTGTTGTGCAATCAACGTTATTTGATATC 1413
DB 1645 GAGGTGTGACAAAGATTTCAAGACCTTTGTTGTGCAATCAACGTTATTTGATATC 1704

QY 1414 AATGATCAGATCCCATCTTTGAAAAATCAGATTATGAACCTGACTCTTGCTGAAGAC 1473
DB 1705 AATGATCAGATCCCATCTTTGAAAAATCAGATTATGAACCTGACTCTTGCTGAAGAC 1764
QY 1474 ACAACATTTGGGTCCACATCTTAACATTCAGGCCATCTGATGAGCCATTTACT 1533
DB 1765 ACAACATTTGGGTCCACATCTTAACATTCAGGCCATCTGATGAGCCATTTACT 1824
QY 1534 GGGAGTTCTAAATTTCTGATATCAATCAATTAAGGAGACATGAGGAGCCCTGGGGGT 1593
DB 1825 GGGAGTTCTAAATTTCTGATATCAATCAATTAAGGAGACATGAGGAGCCCTGGGGGT 1884
QY 1594 GACACAGATCCCATCAACACCGATATGTATTAATTAAGGCTCTTGAATTTGAA 1653
DB 1885 GACACAGATCCCATCAACACCGATATGTATTAATTAAGGCTCTTGAATTTGAA 1944
QY 1654 ACAGCAGCTGTTTCCAAATTTGTTCAAGAGAAATTCCTGAGCCTTGAATTTGT 1713
DB 1945 ACAGCAGCTGTTTCCAAATTTGTTCAAGAGAAATTCCTGAGCCTTGAATTTGT 2004
QY 1714 GTGAGTACAAATGCAATGTTTTCGCAAGTTCAAGCTTATGTCACAGTGTGAATGA 1773
DB 2005 GTGAGTACAAATGCAATGTTTTCGCAAGTTCAAGCTTATGTCACAGTGTGAATGA 2064
QY 1774 GCACTCAATTTTCCCAACAGTATTCGAAGCGAAAGTCAAGTGAATGATAGGC 1833
DB 2065 GCACTCAATTTTCCCAACAGTATTCGAAGCGAAAGTCAAGTGAATGATAGGC 2124
QY 1834 ACTAAGTGGCAATGATGATGCAAGATTCGAAGAGTCTGGAATTAAGTATTCAGT 1893
DB 2125 ACTAAGTGGCAATGATGATGCAAGATTCGAAGAGTCTGGAATTAAGTATTCAGT 2184
QY 1894 AGGGAGACACAAGAGTGGCTTAATTAATGACACGTGATCTGATGATCTTAACTGTC 1953
DB 2185 AGGGAGACACAAGAGTGGCTTAATTAATGACACGTGATCTGATGATCTTAACTGTC 2244
QY 1954 GCTCCATTTGACAGAGAACCCGAGTCCATATCGGGTCAAGTGGTGGCCACAGAGTA 2013
DB 2245 GCTCCATTTGACAGAGAACCCGAGTCCATATCGGGTCAAGTGGTGGCCACAGAGTA 2304
QY 2014 GGGGGGTCTTCTTAAGCTCTGTGTCAGATTCACCTGATCTTATGAGTGAATGAC 2073
DB 2305 GGGGGGTCTTCTTAAGCTCTGTGTCAGATTCACCTGATCTTATGAGTGAATGAC 2364
QY 2074 AACCTCCCAAGCTACAGGACATCAACAGGCTGTTCTTCCATCTCCCTCAGTGA 2133
DB 2365 AACCTCCCAAGCTACAGGACATCAACAGGCTGTTCTTCCATCTCCCTCAGTGA 2424
QY 2134 CTTGAAAGTCTAATTTGAGGCTACTGATGATGATCAGACCTTAATTTGGGGTCCCAT 2193
DB 2425 CTTGAAAGTCTAATTTGAGGCTACTGATGATGATCAGACCTTAATTTGGGGTCCCAT 2484
QY 2194 TTTACATTTTCCCTGGCAGTGAAGCTTACAAACGACTGGGAAGTTTCCAAATCAAT 2253
DB 2485 TTTACATTTTCCCTGGCAGTGAAGCTTACAAACGACTGGGAAGTTTCCAAATCAAT 2544
QY 2254 GGTATCTATGCCCCGACTGTCTACAGGCAACAGACTTTGAGAGAGGCGTATGTCTC 2313
DB 2545 GGTATCTATGCCCCGACTGTCTACAGGCAACAGACTTTGAGAGAGGCGTATGTCTC 2604
QY 2314 TTGATCCGATCAATGATGAGGGGTGGGCCACCTTGAAGGATTTCTTACCAAGT 2373
DB 2605 TTGATCCGATCAATGATGAGGGGTGGGCCACCTTGAAGGATTTCTTACCAAGT 2664
QY 2374 ACATCTGAGTGTGTGAGAGAGAGTGTTCGCGCAGCAGTCAACAGACTGGGATA 2433
DB 2665 ACATCTGAGTGTGTGAGAGAGAGTGTTCGCGCAGCAGTCAACAGACTGGGATA 2724
QY 2434 CCCACTGTGGGCAATGGCACTGTGATATCTGTCACCACTTCTGCTGATTTGATTAAT 2493
DB 2725 CCCACTGTGGGCAATGGCACTGTGATATCTGTCACCACTTCTGCTGATTTGATTAAT 2784
QY 2494 TTAGCAGTTGTGTTATCCGATTAAGAGATTAAGGAGATTAAGTGTGCT 2553

CURRENT APPLICATION NUMBER: US/09/878.178
CURRENT FILING DATE: 2001-06-08
NUMBER OF SEQ ID NOS: 2237
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1019
LENGTH: 520
TYPE: DNA
ORGANISM: Homo sapiens
US-09-878-178-1019

Query Match 15.4%; Score 514.2; DB 9; Length 520;
Best Local Similarity 99.4%; Pred. No. 1,7e-133;
Matches 516; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2257 ACTCATCCCGACGTGTACAGGACACAGACTTTGAGGAGGGGTATGTCCTTG 2316
DB 1 ACTCATCCCGACGTGTACAGGACACAGACTTTGAGGAGGGGTATGTCCTTG 60
QY 2317 ATCCGATCAATATGAGGGGTGCGCCACCCCTTGAGAGCATTTGTTCTTTACAGTTACA 2376
DB 61 ATCCGATCAATATGAGGGGTGCGCCACCCCTTGAGAGCATTTGTTCTTTACAGTTACA 120
QY 2377 TTCTGCAGTTGTGTGGAAGAAAGTTGTTTCGCGCAGAGGTCAACAGCTGGGATACC 2436
DB 121 TTCTGCAGTTGTGTGGAAGAAAGTTGTTTCGCGCAGAGGTCAACAGCTGGGATACC 180
QY 2437 ACTGTGGGCAATGGCAGTTGATATCTGCTGACCAACCCTTGCGTATGATTAATTTA 2496
DB 181 ACTGTGGGCAATGGCAGTTGATATCTGCTGACCAACCCTTGCGTATGATTAATTTA 240
QY 2497 GCAGTTGTGTTATCCGCATTAAGAAAGATTAAGCAAGATTAATGTTGAAGTCTCAA 2556
DB 241 GCAGTTGTGTTATCCGCATTAAGAAAGATTAAGCAAGATTAATGTTGAAGTCTCAA 300
QY 2557 GCATCTGAAGTCAAACTCTGAGAAAGTGAATTTGAAAAGATTTGAAATTTATAG 2616
DB 301 GCATCTGAAGTCAAACTCTGAGAAAGTGAATTTGAAAAGATTTGAAATTTATAG 360
QY 2617 CAAGTGTATTTAGCAACAACCATCTCATCTTAATCTTAATCTTAATCTTAATCTTAAT 2676
DB 361 CAAGTGTATTTAGCAACAACCATCTCATCTTAATCTTAATCTTAATCTTAATCTTAAT 420
QY 2677 ATTTTAAACAGATATTCCTCTTGTCTTAATATTTGCTAAATTTCTTTTGGAG 2736
DB 421 ATTTTAAACAGATATTCCTCTTGTCTTAATATTTGCTAAATTTCTTTTGGAG 480
QY 2737 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 2775
DB 481 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 519

RESULT 14
US-10-146-502-1019
Sequence 1019, Application US/10146502
Publication No. US20030069180A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Harlocker, Susan L.
APPLICANT: Secretist, Heather
APPLICANT: Wang, Aljun
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.527C2
CURRENT APPLICATION NUMBER: US/10/146.502
CURRENT FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2241
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1019
LENGTH: 520
TYPE: DNA
ORGANISM: Homo sapiens
US-10-146-502-1019

Query Match 15.4%; Score 514.2; DB 9; Length 520;
Best Local Similarity 99.4%; Pred. No. 1,7e-133;
Matches 516; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2257 ACTCATCCCGACGTGTACAGGACACAGACTTTGAGGAGGGGTATGTCCTTG 2316
DB 1 ACTCATCCCGACGTGTACAGGACACAGACTTTGAGGAGGGGTATGTCCTTG 60
QY 2317 ATCCGATCAATATGAGGGGTGCGCCACCCCTTGAGAGCATTTGTTCTTTACAGTTACA 2376
DB 61 ATCCGATCAATATGAGGGGTGCGCCACCCCTTGAGAGCATTTGTTCTTTACAGTTACA 120
QY 2377 TTCTGCAGTTGTGTGGAAGAAAGTTGTTTCGCGCAGAGGTCAACAGCTGGGATACC 2436
DB 121 TTCTGCAGTTGTGTGGAAGAAAGTTGTTTCGCGCAGAGGTCAACAGCTGGGATACC 180
QY 2437 ACTGTGGGCAATGGCAGTTGATATCTGCTGACCAACCCTTGCGTATGATTAATTTA 2496
DB 181 ACTGTGGGCAATGGCAGTTGATATCTGCTGACCAACCCTTGCGTATGATTAATTTA 240
QY 2497 GCAGTTGTGTTATCCGCATTAAGAAAGATTAAGCAAGATTAATGTTGAAGTCTCAA 2556
DB 241 GCAGTTGTGTTATCCGCATTAAGAAAGATTAAGCAAGATTAATGTTGAAGTCTCAA 300
QY 2557 GCATCTGAAGTCAAACTCTGAGAAAGTGAATTTGAAAAGATTTGAAATTTATAG 2616
DB 301 GCATCTGAAGTCAAACTCTGAGAAAGTGAATTTGAAAAGATTTGAAATTTATAG 360
QY 2617 CAAGTGTATTTAGCAACAACCATCTCATCTTAATCTTAATCTTAATCTTAATCTTAAT 2676
DB 361 CAAGTGTATTTAGCAACAACCATCTCATCTTAATCTTAATCTTAATCTTAATCTTAAT 420
QY 2677 ATTTTAAACAGATATTCCTCTTGTCTTAATATTTGCTAAATTTCTTTTGGAG 2736
DB 421 ATTTTAAACAGATATTCCTCTTGTCTTAATATTTGCTAAATTTCTTTTGGAG 480
QY 2737 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 2775
DB 481 GTGAGTCTTGTCTGTGCGCCAGGCTGAGATACAGTGG 519

RESULT 15
US-10-066-543-3247
Sequence 3247, Application US/10066543
Publication No. US20030087818A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Pyle, Ruth A.
APPLICANT: Xu, Jiangchun
APPLICANT: Indira, Carol Joseph
APPLICANT: Lodes, Michael J.
APPLICANT: Secretist, Heather
APPLICANT: Carter, Darick
APPLICANT: Fanger, Gary R.
APPLICANT: Smith, Carole L.
APPLICANT: Durham, Margarita
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.563
CURRENT APPLICATION NUMBER: US/10/066.543
CURRENT FILING DATE: 2002-01-31
NUMBER OF SEQ ID NOS: 3417
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3247
LENGTH: 517
TYPE: DNA
ORGANISM: Homo sapiens
US-10-066-543-3247

Query Match 15.3%; Score 512.4; DB 9; Length 517;
Best Local Similarity 99.8%; Pred. No. 5.4e-133;

Matches 513; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 2257 ACTCATGCCGACCTGTCTACACGACACAGACTTTGAGAGAGGCGGTATGTCGTCTG 2316
Db 1 ACTCATGCCGACCTGTCTACACGACACAGACTTTGAGAGAGGCGGTATGTCGTCTG 60
QY 2317 ATCCGATCAATGATGGGGGTGGCCACCCCTTGGAGGCAATGTTCTTTACCAATTACA 2376
Db 61 ATCCGATCAATGATGGGGGTGGCCACCCCTTGGAGGCAATGTTCTTTACCAATTACA 120
QY 2377 TTCTGCAATTGTGTGAAGAAATTTTCCGCCACAGAGGTCAACAGACTGGGATACCC 2436
Db 121 TTCTGCAATTGTGTGAAGAAATTTTCCGCCACAGAGGTCAACAGACTGGGATACCC 180
QY 2437 ACTGTGGGCATGGAGTTGGTATCTGCTGACCAACCCCTTCTGGTATGGTATATTTTA 2496
Db 181 ACTGTGGGCATGGAGTTGGTATCTGCTGACCAACCCCTTCTGGTATGGTATATTTTA 240
QY 2497 GCAATTTGTTTATCCGATTAAGAGGATTAAGGCAAAAGATATGTTGAAAGTCTCAA 2556
Db .241 GCAATTTGTTTATCCGATTAAGAGGATTAAGGCAAAAGATATGTTGAAAGTCTCAA 300
QY 2557 GCATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAAGAAATGTTGAATTTATATAG 2616
Db 301 GCATCTGAAGTCAAACTCTGAGAGCTGAATTTGAAAAGAAATGTTGAATTTATATAG 360
QY 2617 CAAGTGTATTTTCGCAACCAACCATTCATTCCTTACTTTTCATCACTGATGATATA 2676
Db 361 CAAGTGTATTTTCGCAACCAACCATTCATTCCTTACTTTTCATCACTGATGATATA 420
QY 2677 ATTTTAAACAGATATTTCCCTCTGTCCTTAAATATTTGCTAAATATTTCTTTTGGAG 2736
Db 421 ATTTTAAACAGATATTTCCCTCTGTCCTTAAATATTTGCTAAATATTTCTTTTGGAG 480
QY 2737 GTGAGTCTTGCTCTGTGCGCCAGGCTGAGATAC 2770
Db 481 GTGAGTCTTGCTCTGTGCGCCAGGCTGAGATAC 514
```

Search completed: June 22, 2003, 01:03:24
Job time : 453 secs

1	719.4	21.5	1125	14	BM918123	EM918123	AGENCOURT
2	609.8	18.2	613	10	AW859603	MRI-CT032	
3	594.8	17.8	625	12	BE8596136	BE8596136	
4	540.8	16.2	573	9	A1347909	A1347909	GP06004.x
5	524.8	15.7	537	9	A1732453	A1732453	z190F07.x
6	516.8	15.4	540	13	BG378853	BG378853	PM1-CN00053

7	513.6	15.4	590	10	AM859530	MR1-CT035
8	500.4	15.0	519	10	AM859550	MR1-CT035
9	500.2	15.0	514	9	AT1733765	AT1733765 z190f07.y
10	471.2	14.1	563	10	AM862312	RC4-CT036
11	470.4	14.1	475	9	AT1626303	AT1626303 qk41h05.x
12	468.4	14.0	530	9	AA1023326	AA1023326 z190f07.x
13	467.2	14.0	472	14	BM799332	K-EST0083
14	464.2	13.9	486	13	BG981361	CM4-CN008
15	452.8	13.5	477	12	BF334481	MR1-CT035
16	451.6	13.5	495	10	AM853021	RC1-CT024
17	449.6	13.4	554	9	AA088861	z190f07.s
18	440.4	13.4	457	9	AT920859	w015b06.x
19	435.2	13.0	706	10	AM735947	RC0-CT020
20	435	13.0	455	12	BE656126	BE656126 MR1-CT025
21	434.6	13.0	873	12	BF582721	BF582721 602094137
22	430.8	12.9	620	10	AM737543	QVO-CT022
23	418.4	12.5	507	10	AM753130	RC1-CT024
24	410.4	12.3	464	10	AM731746	CM2-ST022
25	408.4	12.2	428	9	AT623483	ts19c10.x
26	404.2	12.1	436	13	BG981357	CM4-CN008
27	401.8	12.0	406	10	AM735184	RC1-CT019
28	401.8	12.0	408	9	AT1922856	wo14b06.x
29	401.6	12.0	551	10	AM887172	RC1-OT008
30	399.4	11.9	461	10	AM859489	MR1-CT035
31	394.6	11.8	657	12	BF540425	BF5840425 602096265
32	394.4	11.8	405	13	EG980846	CM4-CN008
33	391	11.7	403	9	AA053188	AA053188 z172a06.x
34	387.4	11.6	551	12	BE808624	BE808624 z13876.MA
35	383.8	11.5	441	14	BQ330693	RC4-CT084
36	383.8	11.5	771	9	AT1566493	tr60e08.x
37	383.2	11.5	753	12	BF577684	BF577684 602092567
38	377.8	11.3	435	13	BG981295	CM4-CN008
39	375.4	11.2	417	10	AM7351659	QV2-CT014
40	372.4	11.1	408	14	BQ330866	RC4-CT082
41	369	11.0	408	12	BF754330	IL5-CT051
42	368.6	11.0	477	10	AM859516	MR1-CT035
43	364.4	10.9	395	10	AM780504	RC3-CN001
44	364.2	10.9	447	10	AM859567	MR1-CT035
45	360.4	10.8	454	9	AA3053406	AA3053406 EST176656

sou

```

Plate: L1CM2016  row: h  column: 10
High quality sequence stop: 662.
Location/Qualifiers
FEATURES
    source          1..1125

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="IMAGE:5485497"
/tissue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: blood; Vector: pORF1; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
SuperScript II RT (Life Technologies). Note: this is a
NIH-MGC Library."
BASE COUNT      259 a      321 c      258 g      286 t      1 others
ORIGIN
Query Match      21.5%; Score 719.4; DB 14; Length 1125;
Best Local Similarity 97.0%; Pred. No. 2.6e-128;
Matches 754; Conservative 0; Mismatches 21; Indels 2; Gaps 2;

QY      1825 GCTATAGGCACTAAGTGGGCAATGTGACTGCCAAGATCCAGAAGTCTTGACATTAAGC 1884
      1 GCTATAGGCACTAAGTGGGCAATGTGACTGCCAAGATCCAGAAGTCTTGACATTAAGC 60
DB
QY      1885 TATTCACCTGAGGGGAGACACAAAGGTTGGCTTAAATGACACGCTGACCTGTGAGATC 1944
      61 TATTCACCTGAGGGGAGACACAAAGGTTGGCTTAAATGACACGCTGACCTGTGAGATC 120
DB
QY      1945 TTATGATGCTCCATTTGACAGAGAACCCGGAAGTCCATTCGGGATCAAGTGGTGGCC 2004
      121 TTATGATGCTCCATTTGACAGAGAACCCGGAAGTCCATTCGGGATCAAGTGGTGGCC 180
DB
QY      2005 ACAGAGATGAGGGGGGCTCTTCTTAAGCTGTGTGACAGTTCACTGATCTTATGAT 2064
      181 ACAGAGATGAGGGGGGCTCTTCTTAAGCTGTGTGACAGTTCACTGATCTTATGAT 240
DB
QY      2065 GTGAATGACAACTCCCAAGCTAGCCAGACATCACGGGCTGTTCTTGTCCATCCC 2124
      241 GTGAATGACAACTCCCAAGCTAGCCAGACATCACGGGCTGTTCTTGTCCATCCC 300
DB
QY      2125 CTCAGTGAACCTGGAAGTCTCATTTTGGAGGCTACTGATGATGATGATGATGATGAT 2184
      301 CTCAGTGAACCTGGAAGTCTCATTTTGGAGGCTACTGATGATGATGATGATGATGAT 360
DB
QY      2185 GGTCCCATTTTACATTTTCCCTCGGAGTGAAGCTTACAAAGCACTGGAAGTTCC 2244
      361 GGTCCCATTTTACATTTTCCCTCGGAGTGAAGCTTACAAAGCACTGGAAGTTCC 420
DB
QY      2245 AAAATCAATGATCTACTGCCCCGACTGCTACAGGACACAGACTTTGAGAGAGAGGCG 2304
      421 AAAATCAATGATCTACTGCCCCGACTGCTACAGGACACAGACTTTGAGAGAGAGGCG 480
DB
QY      2305 TATGTCGTCGTGATCCGCATCAATGATGAGGGGTGGCCACCTTGGAAGGATTTCT 2364
      481 TATGTCGTCGTGATCCGCATCAATGATGAGGGGTGGCCACCTTGGAAGGATTTCT 540
DB
QY      2365 TTACAGTTCATTTCTGCAAGTTGTGTGAGAGAAAGTTGTTTCCGGCCACAGGTCCACG 2424
      541 TTACAGTTCATTTCTGCAAGTTGTGTGAGAGAAAGTTGTTTCCGGCCACAGGTCCACG 600
DB
QY      2425 ACTGGGATACCCCATCTGAGGAGATGGGAGTGTATACCTGACACACCTTCTGTGTAT 2484
      601 ACTGGGATACCCCATCTGAGGAGATGGGAGTGTATACCTGACACACCTTCTGTGTAT 660
DB
QY      2485 GGTATTAATTTAG-CAGTTGTATTATCCGCATTAAGAGAGATTAAGGCAAGATTAATGT 2543
      661 GGTATTAATTTAG-CAGTTGTATTATCCGCATTAAGAGAGATTAAGGCAAGATTAATGT 720
DB
QY      2544 TGAAGTGTCTCAAGCATC-TGAAGTCAAACTCTGAGAGTGAATTTGAAAGAGAA 2599
      721 TGAAGTGTCTCAAGCATC-TGAAGTCAAACTCTGAGAGTGAATTTGAAAGAGAA 777
DB

```

```

RESULT 2
AM859603
LOCUS
DEFINITION
M81-CT0355-180200-007-e07 CT0355 Homo sapiens cDNA, mRNA sequence.
ACCESSION
AM859603.1 GI:7955296
VERSION
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 613)
REFERENCE
Dias Neto, E., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Spotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663.
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=M81-CT0355-180
200-007-e07&t3=2000-02-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 613.
FEATURES
source
1..613
location/Qualifiers
1..613
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0355"
/def_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      195 a      125 c      134 g      159 t
ORIGIN
Query Match      18.2%; Score 609.8; DB 10; Length 613;
Best Local Similarity 99.7%; Pred. No. 3.4e-107;
Matches 611; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1318 AAGAGCAGAGACTCTCTCACTGATCAACTTACAGATGAGGTGCTGACAAAGTTTCAAG 1377
      1 AAGAGCAGAGACTCTCTCACTGATCAACTTACAGATGAGGTGCTGACAAAGTTTCAAG 60
DB
QY      1378 ACCCTTTGTTTGTGCAAAACAAGTATGATGATCAATGATGATGATGATGATGATGAT 1437
      61 ACCCTTTGTTTGTGCAAAACAAGTATGATGATGATGATGATGATGATGATGATGAT 120
DB
QY      1438 AAATCAGATTAATGAAACCTGACTCTGTGAGAGACAAACATTTGGTCCACATCTTA 1497
      121 AAATCAGATTAATGAAACCTGACTCTGTGAGAGACAAACATTTGGTCCACATCTTA 180
DB
QY      1498 ACCATCAGGCCCATGATGCTGATGAGCCATTTTCTGGAGTTCTTAAATTCGTATCAT 1557
      181 ACCATCAGGCCCATGATGCTGATGAGCCATTTTCTGGAGTTCTTAAATTCGTATCAT 240
DB

```

QY 1558 ATCAATAAGGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACACC 1617
 DB 241 ATCAATAAGGAGAGACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACACC 300
 QY 1618 GGAATATGTCATTAATTAAGGCGCTGATTTTGAACACAGAGTGTTCACATATGTC 1677
 DB 301 GGAATATGTCATTAATTAAGGCGCTGATTTTGAACACAGAGTGTTCACATATGTC 360
 QY 1678 TTCAAGCAGAAAATCTTGAGCCCTTGAAGTGTGTTGTTGTTGTTGTTGTTGTTGTT 1737
 DB 361 TTCAAGCAGAAAATCTTGAGCCCTTGAAGTGTGTTGTTGTTGTTGTTGTTGTTGTT 420
 QY 1738 GCCAAGTTCACGCTTATTTGACAGATGTGAATGAGCACCTCAATTTTCCCAACGTA 1797
 DB 421 GCCAAGTTCACGCTTATTTGACAGATGTGAATGAGCACCTCAATTTTCCCAACGTA 480
 QY 1798 TTCAAGCAGAAAATCTTGAGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1857
 DB 481 TTCAAGCAGAAAATCTTGAGGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540
 QY 1858 AAGATTCAGAAAGTCTGACATAGCTATTTCATCTGAGGGAGACACAGAGGTTGCTT 1917
 DB 541 AAGATTCAGAAAGTCTGACATAGCTATTTCATCTGAGGGAGACACAGAGGTTGCTT 600
 QY 1918 AAAATTGACCAGC 1930
 DB 601 AAAATTGACCAGC 613

RESULT 3
 BE96136/c 625 bp mRNA linear EST 11-SEP-2000
 LOCUS BE96136
 DEFINITION M1-CT0258-040700-002-e10 CT0258 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE96136
 VERSION BE96136.1 GI:10083296
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 625)

REFERENCE
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldan,M.G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 CONTACT: Simpson A.J.G.
 20202663
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

FEATURES
 source
 1. .625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0258"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
 SmaI; A mini-library was made by cloning products derived
 High quality sequence stop: 620.
 Location/Qualifiers

from ORSTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 143 a 141 c 153 g 188 t
 ORIGIN

Query Match 17.8%; Score 594.8; DB 12; Length 625;
 Best Local Similarity 98.8%; Pred. No. 2.6e-104;
 Matches 599; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 277 GGGAGACAGACAAATTTTGTGATGAAACGGAGGAGCTTGTATTACAACAGAGCC 336
 DB 606 GGGAGACAGACAAATTTTGTGATGAAACGGAGGAGCTTGTATTACAACAGAGCC 517
 QY 337 TTGACAGAGAAACAAGATCTACTCAATCTCCAGGTTGCAGCCCTGAGCTAATGGA 336
 DB 546 TTGACAGAGAAACAAGATCTACTCAATCTCCAGGTTGCAGCCCTGAGCTAATGGA 487
 QY 397 ATTATAGTGAAGGTCAGTCCCTATCACCATTAAGTGAAGGACATCAACGACATTCGA 456
 DB 486 ATTATAGTGAAGGTCAGTCCCTATCACCATTAAGTGAAGGACATCAACGACATTCGA 427
 QY 457 CCCACGTTTCTCCAGTCAAGTACGAAAGGCTCAGTAAGGACAGAACTCTGCGCCAGGAAG 516
 DB 426 CCCACGTTTCTCCAGTCAAGTACGAAAGGCTCAGTAAGGACAGAACTCTGCGCCAGGAAG 367
 QY 517 CCTTCTTTGTATGTCAATGCCACAGACTGTGATTCGGCCACTCCCAATGGCCAGCTT 576
 DB 366 CCTTCTTTGTATGTCAATGCCACAGACTGTGATTCGGCCACTCCCAATGGCCAGCTT 307
 QY 577 TATTACAGATTGTCAATCCAGCTTCCAGATCAACAATGTACTTTTCAGATCAAC 636
 DB 306 TATTACAGATTGTCAATCCAGCTTCCAGATCAACAATGTACTTTTCAGATCAAC 247
 QY 637 AACCAAAACGGAGCCATCTCTCTTACCCGAGAGGATCTCAGAAATTTGATCTCTGTAAG 696
 DB 246 AACCAAAACGGAGCCATCTCTCTTACCCGAGAGGATCTCAGAAATTTGATCTCTGTAAG 187
 QY 697 AATCTTCTTATATTTGTTGATCTTCAAGTGAAGACATGGAGGCCAGAGTGAATTC 756
 DB 186 AATCTTCTTATATTTGTTGATCTTCAAGTGAAGACATGGAGGCCAGAGTGAATTC 127
 QY 757 TTCAGGATATACCAATCTGTGATATCAATGAGACAGAAATTTTGAAGACCAAAA 816
 DB 126 TTCAGGATATACCAATCTGTGATATCAATGAGACAGAAATTTTGAAGACCAAAA 67
 QY 817 CCTGTGAGATGGTGAAGAACTCAACTGATCTCAACCCCATCAAAATCACTAGTGGCG 876
 DB 66 CCTGTGAGATGGTGAAGAACTCAACTGATCTCAACCCCATCAAAATCACTAGTGGCG 7
 QY 877 TGGAAAT 882
 DB 6 TGGAAAT 1

RESULT 4
 A1347909/c 573 bp mRNA linear EST 30-DEC-1998
 LOCUS A1347909
 DEFINITION gp060e04.x1 NCI CGAP C08 Homo sapiens cDNA clone IMAGE:1927422.3,
 similar to TR_015336 Q15336 LI-CADHERIN. [1], mRNA sequence.
 ACCESSION A1347909
 VERSION A1347909.1 GI:4085115
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 573)
 REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

Db 297 ATTATTCATGACATTTTCTCTCTGCAATGAGCTAGCTTGTGTTTCCCTTTT 228
 QY 3106 GGGGCAAGACAGACTCAATTAATTTCTGATCATTTTCTTTATCAAGAGATATATCA 3165
 Db 237 GGGGCAAGACAGACTCAATTAATTTCTGATCATTTTCTTTATCAAGAGATATATCA 178
 QY 3166 GTGTTGTCATAGAACCTGCTGGATTCATTATGTTTTTCTGATTCATCTGTGTC 3225
 Db 177 GTGTTGTCATAGAACCTGCTGGATTCATTATGTTTTTCTGATTCATCTGTGTC 118
 QY 3226 CCCTTATCCTTGACTCTCTTGTGATTTTCACTGAATTTCAAACTTTGTCAAGAAATAA 3285
 Db 117 CCCTTATCCTTGACTCTCTTGTGATTTTCACTGAATTTCAAACTTTGTCAAGAAATAA 58
 QY 3286 AAAGTGAGACTCAGAAAAATAATAATAAAGAACAGCCTTTTGC 3333
 Db 57 AACGTGAGACTCAGAAAAATAATAATAAAGAACAGCCTTTTCC 10

RESULT 6
 BG978653/c 540 bp mRNA linear EST 12-JUN-2001
 LOCUS PM1-CN0098-020101-004-e02 CN0098 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BG978653
 ACCESSION BG978653.1 GI:14381388
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 540)

REFERENCE
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

TITLE
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM1&fz=PM1-CN0098-
 020101-004-e02&fz=2001-01-02&fz=1)
 Seq primer: puc 18 forward
 High quality sequence start: 7
 High quality sequence stop: 538.
 Location/Qualifiers

FEATURES
 source
 1..540
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CN0098"
 /dev_stage="Adult"
 /note="Organ: colon normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent Application
 No.196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 136 a 136 c 127 g 140 t 1 others
 Query Match 15.4%; Score 516.8; DB 13; Length 540;

Best Local Similarity 98.3%; Pred. No. 2,6e-89;
 Matches 532; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1754 TTGTGACAGATGTAAGTGAAGACACCTCAATTTTCCCAACGATTTCCAGCGAAATGCA 1813
 Db 540 TTGTGACAGATGTAAGTGAAGACACCTCAATTTTCCCAACGATTTCCAGCGAAATGCA 481
 QY 1814 GTGAGATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1873
 Db 480 GTGAGATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 421
 QY 1874 TGAACATTAAGTATTTCACTGAGGGGAGACACAAGAGTTGGCTTAAATTAATGACCACTGA 1933
 Db 420 TGAACATTAAGTATTTCACTGAGGGGAGACACAAGAGTTGGCTTAAATTAATGACCACTGA 361
 QY 1934 CTGGTGAAGTCTTAATGATGCTCCATTTGAGACAGAAAGCCGATTCATATCCAGGTAC 1993
 Db 360 CTGGTGAAGTCTTAATGATGCTCCATTTGAGACAGAAAGCCGATTCATATCCAGGTAC 301
 QY 1994 AAGTGTGAGCCACAGAAAGTGAAGGGGGCTTCTTCTAGCTCTGTGACAGATTCACCTGA 2053
 Db 300 AAGTGTGAGCCACAGAAAGTGAAGGGGGCTTCTTCTAGCTCTGTGACAGATTCACCTGA 241
 QY 2054 TCGTTATGATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2113
 Db 240 TCGTTATGATGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 181
 QY 2114 TCTGCAATCCCTCATGTCAGTCCGAGTCAATTTTCCAGGCTACTGATGATGATCAGC 2173
 Db 180 TCTGCAATCCCTCATGTCAGTCCGAGTCAATTTTCCAGGCTACTGATGATGATCAGC 121
 QY 2174 ACTTATTTGGGGGTGCCATTTTATCATTTTCCCTGCGAGTGAAGTGAAGTGAAGTGA 2233
 Db 120 ACTTATTTGGGGGTGCCATTTTATCATTTTCCCTGCGAGTGAAGTGAAGTGAAGTGAAGTGA 61
 QY 2234 GGGAGTTTCCAAATATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2293
 Db 60 GGGAGTTTCCAAATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2
 QY 2294 A 2294
 Db 1 A 1

RESULT 7
 AM859530 590 bp mRNA linear EST 19-MAY-2000
 LOCUS MR1-CT0355-160200-003-e05 CT0355 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM859530
 ACCESSION AM859530.1 GI:7955223
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 590)

REFERENCE
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?fl=et2=MR1-CT0355-160>)
 200-003-e05et3=2000-02-16et4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 22
 High quality sequence stop: 577.
 Location/Qualifiers

FEATURES
 source
 1. 590
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="CT0355"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 183 a 127 c 132 g 148 t
 ORIGIN

Query Match 15.4%; Score 513.6; DB 10; Length 590;
 Best Local Similarity 99.1%; Pred. No. 1e-88; 4; Indels 1; Gaps 1;
 Matches 527; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1330 ACTCCCTAGTACAACTTAACGATAGAGGTCTGACAAAGATTTCAGACCTTTGTTT 1389
 DB 19 ACGCTCAGTACAACTTAACGATAGAGGTCTGACAAAG-TCAGACCTTTGTTT 77
 QY 1390 GTGCAATCAACGTTATTTGATATCAATGATAGATCCCATCTTTGAAAAATCGATTAT 1449
 DB 78 GTGCAATCAACGTTATTTGATATCAATGATAGATCCCATCTTTGAAAAATCGATTAT 137
 QY 1450 GGAACCTGACTCTTGTGTAAGACACAAACCTTGGGTCCACATCTTAACCATCCAGCC 1509
 DB 138 GGAACCTGACTCTTGTGTAAGACACAAACCTTGGGTCCACATCTTAACCATCCAGCC 197
 QY 1510 ACTGATCTGATGAGCACTTACTGGAAGTTCTAAATTCGTATCATATCATTAAGGA 1569
 DB 198 ACTGATCTGATGAGCACTTACTGGAAGTTCTAAATTCGTATCATATCATTAAGGA 257
 QY 1570 GACAGTGAAGGAGCCCTGGGGGTGACACAGATCCCATACCAACCGGATATGTATA 1629
 DB 258 GACAGTGAAGGAGCCCTGGGGGTGACACAGATCCCATACCAACCGGATATGTATA 317
 QY 1630 ATTAAAAAGCCTCTTGTGTAAGACAGAGCTGTTCCAACTTTGTTCAAGCAGAA 1689
 DB 318 ATTAAAAAGCCTCTTGTGTAAGACAGAGCTGTTCCAACTTTGTTCAAGCAGAA 377
 QY 1690 AATCCTGAGCCTTAGTGTGTTGTTGTAAGTACAAATGCAAGTCTTTTGCCAACTTACG 1749
 DB 378 AATCCTGAGCCTTAGTGTGTTGTTGTAAGTACAAATGCAAGTCTTTTGCCAACTTACG 437
 QY 1750 CTTATTTGACAGATGTAAGTGAAGCACTCAATTTTCCCAACGATATTCAGACGAAA 1809
 DB 438 CTTATTTGACAGATGTAAGTGAAGCACTCAATTTTCCCAACGATATTCAGACGAAA 497
 QY 1810 GTCAAGTGAAGTATGAGCACTTAAGTGGGCAATGATGCAAGTGCAGAG 1861
 DB 498 GTCAAGTGAAGTATGAGCACTTAAGTGGGCAATGATGCAAGTGCAGAG 549

RESULT 8
 AM859550 519 bp mRNA linear EST 19-MAY-2000
 LOCUS AM859550
 DEFINITION MR1-CT0355-160200-004-b02 CT0355 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM859550
 VERSION AM859550.1 GI:7955243
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 519)
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Coe, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?fl=et2=MR1-CT0355-160>)
 200-004-b02et3=2000-02-16et4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 81
 High quality sequence stop: 517.
 Location/Qualifiers

FEATURES
 source

1. 519

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1lb="CT0355"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 160 a 111 c 107 g 141 t
 ORIGIN

Query Match 15.0%; Score 500.4; DB 10; Length 519;
 Best Local Similarity 97.9%; Pred. No. 3.7e-86;
 Matches 507; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1329 TACTCCGATACAACTTAACGATAGAGGTCTGACAAAGATTTCAGACCTTTGTTT 1388
 DB 2 TACTCCGATACAACTTAACGATAGAGGTCTGACAAAGATTTCAGACCTTTGTTT 61
 QY 1389 TGTGCAATCAACGTTATTTGATATCAATGATAGATCCCATCTTTGAAAAATCGATTAT 1448
 DB 62 TGTGCAATCAACGTTATTTGATATCAATGATAGATCCCATCTTTGAAAAATCGATTAT 121
 QY 1449 TGGAAACCTGACTCTTGTGTAAGACACAAACATTTGGGTCCACATCTTTAACCATCCAGGC 1508
 DB 122 TGGAAACCTGACTCTTGTGTAAGACACAAACATTTGGGTCCACATCTTTAACCATCCAGGC 181
 QY 1509 CACTGATGCTGATGAGCACTTACTGGAAGTTCTAAATTCGTATCATATCATTAAGG 1568
 DB 182 CACTGATGCTGATGAGCACTTACTGGAAGTTCTAAATTCGTATCATATCATTAAGG 241
 QY 1569 AGACAGTGAAGGAGCCCTGGGGGTGACACAGATCCCATACCAACCGGATATGTAT 1628
 DB 242 AGACAGTGAAGGAGCCCTGGGGGTGACACAGATCCCATACCAACCGGATATGTAT 301
 QY 1629 AATTTAAAAAGCCTCTTGTGTTTGAACAGAGCTGTTTCCAACTTTGTTCAAGCAGAA 1688
 DB 302 AATTTAAAAAGCCTCTTGTGTTTGAACAGAGCTGTTTCCAACTTTGTTCAAGCAGAA 361

QY 1689 AATCCGAGGCTCTAGTGTGGTGAAGTACATGCAAGTCTTTGGCAAGTTCAC 1748
 DB 362 AATTCGAGGCTCTAGTGTGGTGAAGTACATGCAAGTCTTTGGCAAGTTCAC 421
 QY 1749 GCTATTGTGACATGTGAATGAAGCACTTCATTTTCCCAACAGGATTCACGGA 1808
 DB 422 GCTATTGTGACATGTGAATGAAGCACTTCATTTTCCCAACAGGATTCACGGA 481
 QY 1809 AGTCAGTGAAGTGTAGCTATAGGCACTTAAGTGGGCA 1846
 DB 482 AGTCAGTGAAGTGTAGCTATAGGCACTTAAGTGGGCA 519
 RESULT 9
 A1733765 514 bp mRNA linear EST 14-JUN-1999
 LOCUS 2190f07.Y5 StrataGene colon (#937204) Homo sapiens cDNA clone
 DEFINITION IMAGE:511909.5, similar to TR:Q12864 Q12864 INTESTINAL
 PEPTIDE-ASSOCIATED TRANSPORTER HPT-1. [1], mRNA sequence.
 A1733765
 VERSION A1733765.1 GI:5054878
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 514)
 REFERENCE NCBI/NIH-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS National Cancer Institute / National Institute of Dental Research,
 TITL Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Other ESTs: z190f07.x5
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-rcmail.nih.gov
 This clone is available royalty-free through LBNL; contact the
 IMAGE Consortium (infoimage.lbnl.gov) for further information.
 This read is a RESEQUENCE of a previously sequenced human clone
 Original clone citation: Washu-Merck EST Project
 This read has been verified (found to hit its original self in the
 correct orientation)
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40RP from Gdbco
 High quality sequence stop: 405.
 Location/Qualifiers
 1..514
 /organism="Homo sapiens"
 /db_xref="GDB:384404"
 /db_xref="taxon:9606"
 /clone="IMAGE:511909"
 /clone_1db="Stratagene colon (#937204)"
 /tissue type="tumor"
 /cell_line="H84 carcinoma cell line"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: colon; Vector: pBluescript SK-; Site_1:
 EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
 Oligo dT. T-84 colonic epithelial cell line. Average
 insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
 sequence: 5' GAATTCGACGAG 3' ~3' adaptor sequence: 5'
 CTCGAGTTTCTTTTCTTTTCTTTT 3'"

BASE COUNT 135 a 115 c 132 g 132 t
 ORIGIN
 Query Match 15.0%; Score 500.2; DB 9; Length 514;
 Best Local Similarity 98.4%; Pred. No. 4e-86; Indels 0; Gaps 0;
 Matches 505; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1677 GTTCAAGCAGAAATCTAGGCTCTAGTGTGGTGAAGTACATGCAAGTCTTT 1736
 DB 1 GTTCAAGCAGAAATCTAGGCTCTAGTGTGGTGAAGTACATGCAAGTCTTT 60
 QY 1737 TGGCAAGTTCACGCTTATTGTGACATGTGAATGAAGCACTTCATTTTCCCAACAGCT 1796
 DB 61 TGGCAAGTTCACGCTTATTGTGACATGTGAATGAAGCACTTCATTTTCCCAACAGCT 120

QY 1797 ATTCACGCAAAAGTCAGTGAAGATGTAGCTATAGGCACTAAAGTGGCAATGTGACTGC 1856
 DB 121 ATTCCAGCGAAAGTCAGTGAAGATGTAGCTATAGGCACTAAAGTGGCAATGTGACTGC 180
 QY 1857 CAAGATTCACGAAGGTCTGACATTAAGCTTATTCCTAGAGGGAGACACAAAGGTTGGCT 1916
 DB 181 CAAGATTCACGAAGGTCTGACATTAAGCTTATTCCTAGAGGGAGACACAAAGGTTGGCT 240
 QY 1917 TAAATTTGACACATGATCTGTGAGATCTTTAGTGTGGCTTCATTGACACAGAAAGCCGG 1976
 DB 241 TAAATTTGACACATGATCTGTGAGATCTTTAGTGTGGCTTCATTGACACAGAAAGCCGG 300
 QY 1977 AAGTCATATCGGGTACAGATGTGGCCACAGAAAGTGGGGGCTCTTCTTAAGCTCTGT 2036
 DB 301 AAGTCATATCGGGTACAGATGTGGCCACAGAAAGTGGGGGCTCTTCTTGAAGCTCTGT 360
 QY 2037 GTCAAGTTCACCTGATATCTTATGTGAATGAATGACAACTCCAGAGCTAGCCAAAGA 2096
 DB 361 GTCAAGTTCACCTGATATCTTATGTGAATGAATGACAACTCCAGAGCTAGCCAAAGA 420
 QY 2097 CTACAGCGGCTTGTCTTCTTCCCATATCCCTCAGTGCACCTGGAAGTCTCATTTTCAGGC 2156
 DB 421 CTACAGCGGCTTGTCTTCTTCCCATATCCCTCAGTGCACCTGGAAGTCTCATTTTCAGGC 480
 QY 2157 TACTGATGATGATCAGACCTTATTTTGGGGTCC 2189
 DB 481 TACTGATGATGATCAGACCTTATTTTGGGATCCC 513
 RESULT 10
 AM862312 563 bp mRNA linear EST 19-MAY-2000
 LOCUS AM862312/c
 DEFINITION R04-CT0367-130200-011-b08 CT0367 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM862312
 VERSION AM862312.1 GI:7958007
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 563)
 REFERENCE Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brites,M.R.,
 Nagai,M.A., da Silva,M.Jr., Borlin,S., Costa,R.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 JOURNAL MEDLINE
 COMMENT 20202663
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=RC4-CT0367-130
 200-011-b08&ct=2000-02-13&rt=1)
 Seq primer: puc 18 forward
 High quality sequence start: 19
 High quality sequence stop: 499.
 Location/Qualifiers
 1..563
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1db="CT0367"
 /dev_stage="Adult"

FEATURES
 source
 1..563
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_1db="CT0367"
 /dev_stage="Adult"

/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 155 a 113 c 121 g 174 t
ORIGIN

Query Match 14.1%; Score 471.2; DB 10; Length 563;
Best Local Similarity 94.2%; Pred. No. 1.5e-80;

Matches 533; Conservative 0; Mismatches 28; Indels 5; Gaps 4;

QY 960 TATTAGGAGTCAAGCCCTTGGACCGAAGAAAGATGCAATGTTTTTATGACGT 1019
DB 563 TATTAGGAGTCAAGCCCTTGGACCGAAGAAAGATGCAATGTTTTTATGACGT 504
QY 1020 TGGAAAGATGAGTACGGAACCACTTTCATCCGCTGGAATTCATGTAAGTTAA 1079
DB 503 TCCAAAGATGAGTACGGAACCACTTTCATCCGCTGGAATTCATGTAAGTTAA 444
QY 1080 AGATATTATGATATCCACTACATGTCCTGCAACAGTAACTGATTGAGTCCAGA 1139
DB 443 AGATATTATGATATCCACTACATGTCCTGCAACAGTAACTGATTGAGTCCAGA 384
QY 1140 GAATGAACGACTGGGTAAACAGTATCGGACCCCTTACTGACATGACAGGATGAGAAA 1199
DB 383 GAATGAACGACTGGGTAAACAGTATCGGACCCCTTACTGACATGACAGGATGAGAAA 324
QY 1200 -TATGCAACAGTTTCTTAACTAAGATTTGAGCAAACTCCCAATCTCCATGG 1258
DB 323 GTACTGCCACAGTTTCTTAACTAAGATTTGAGCAAACTCCCAATCTCCATGG 264
QY 1259 ATGCACTCTTCTTAACTAAGATTTGAGCAAACTCCCAATCTCCATGG 1318
DB 263 ATGCACTCTTCTTAACTAAGATTTGAGCAAACTCCCAATCTCCATGG 204
QY 1319 AGAAGCAAGATATCTCTGATCACTTAAAGATGAGTGTCTGACAAAGATTTCAAGA 1378
DB 203 AGAAGCAAGATATCTCTGATCACTTAAAGATGAGTGTCTGACAAAGATTTCAAGA 144
QY 1379 CCCCTTGTGTCGAATCAAGTTA-TTGATATCAATGATGATCCCATCTTTGAA 1437
DB 143 CCCCTTGTGTCGAATCAAGTTA-TTGATATCAATGATGATCCCATCTTTG-T 85
QY 1438 AATCAGATTATGAAAACCTGACTCTGCTGAGAACCAAACTTGGTCCACATCTTA 1497
DB 84 AATCAGATTATGAAAACCTGACTCTGCTGAGAACCAAACTTGGTCCACATCTTA 27
QY 1498 ACCATCCAGGCCACTGATGCTGATGA 1523
DB 26 ACCATCCAGGCCACTGATGCTGATGA 1

RESULT 11
AI262603/c 475 bp mRNA linear EST 28-JAN-1999

LOCUS AI262603
DEFINITION gk41h05.x1 NCI CGAP C08 Homo sapiens cDNA clone IMAGE:187157 3'
similar to contains Alu repetitive element; mRNA sequence.

ACCESSION AI262603
VERSION AI262603.1 GI:3870806
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 475)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nlm.nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/biopr/image/image.html

Insert Length: 838 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 474.

FEATURES

source

1. 475
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:187157"
/clone_lib="NCI CGAP C08"
/issue_type="adenocarcinoma"
/lab_host="DH10B"

/note="Organ: colon; Vector: pT773D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT773
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."

BASE COUNT 166 a 80 c 107 g 122 t
ORIGIN

Query Match 14.1%; Score 470.4; DB 9; Length 475;
Best Local Similarity 99.8%; Pred. No. 2.2e-80;

Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2862 CCCGACCAACATGCCCAGCTAATTTTGTATTTTAAATAGAGAGGGGTTTGGCATTTGG 2921
DB 475 CCCGACCAACATGCCCAGCTAATTTTGTATTTTAAATAGAGAGGGGTTTGGCATTTGG 416
QY 2922 CAGGCTGCTTGAATCTCTGACGTCAAGTATCTGCTGCTTGGTCTCCCAATACAG 2981
DB 415 CAGGCTGCTTGAATCTCTGACGTCAAGTATCTGCTGCTTGGTCTCCCAATACAG 356
QY 2982 GCATGAACCACTGACCACTTACTAGATTTTCAATGTCATAGACATTAGAGATT 3041
DB 355 GCATGAACCACTGACCACTTACTAGATTTTCAATGTCATAGACATTAGAGATT 296
QY 3042 TTTTCAATTTTCAATGACATTTTCTCTGCAAAATGGCTTACCTAGTTGTTTTGCC 3101
DB 295 TTTTCAATTTTCAATGACATTTTCTCTGCAAAATGGCTTACCTAGTTGTTTTGCC 236
QY 3102 TTTTGGGGCAAGACAGACTATTAATATCTGTACATTTTCTTTAATGAAGATAT 3161
DB 235 TTTTGGGGCAAGACAGACTATTAATATCTGTACATTTTCTTTAATGAAGATAT 176
QY 3162 ATCAGTGTCTCATGAACTGCTGATTCATTTATTTTCTGATTCATCTG 3221
DB 175 ATCAGTGTCTCATGAACTGCTGATTCATTTATTTTCTGATTCATCTG 116
QY 3222 TGTCCCTTCACTCTGACTCTTGGTATTTCACTGAATTTCAACATTTGTGAGAGAA 3281
DB 115 TGTCCCTTCACTCTGACTCTTGGTATTTCACTGAATTTCAACATTTGTGAGAGAA 56
QY 3282 GAAAAAAGTGAAGCTCAGAAAAATTAATTAATTAAGAACGCTTTTGC 3333
DB 55 GAAAAAAGTGAAGCTCAGAAAAATTAATTAATTAAGAACGCTTTTGC 4

RESULT 12

LOCUS AA102326
DEFINITION AA102326 530 bp mRNA linear EST 28-OCT-1996
IMAGE:511909 5' similar to TR:G854175 G854175 LI-CADHERIN. [1];

ACCESSION mRNA sequence.
VERSION AA102326
KEYWORDS AA102326.1 GI:1647016
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 530)
AUTHORS Hallier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chisese, S., Dietrich, N., Dubugue, T., Favello, A., Gish, W., Hawkins,
'M., Hulman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Nardis, E., Moore,
'B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasakis, E.,
Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
TITLE
JOURNAL
MEDLINE
COMMENT 97044478
Contact: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28W13 rev2 from Amersham
High quality sequence stop: 347.
Location/Qualifiers
1..530
/organism="Homo sapiens"
/db_xref="GDB:3844404"
/db_xref="taxon:9606"
/clone="IMAGE:511909"
/clone_lib="Stratagene colon (#937204)"
/tissue_type="tumor"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT: T-84 colonic epithelial cell line. Average
insert size: 1.0 Kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGGACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTT 3'."
BASE COUNT 136 a 116 c 135 g 139 t 4 others
ORIGIN
Query Match 14.0%; Score 468.4; DB 9; Length 530;
Best Local Similarity 97.5%; Pred. No. 5.2e-80;
Matches 517; Conservative 0; Mismatches 8; Indels 5; Gaps 4;
QY 1678 TTCAAGCAGAAATCTGAGCCCTGAGTGTGTTGTGTGAATGCAATGCAAGTCTTTT 1737
DB 1 TTCAAGCAGAAATCTGAGCCCTGAGTGTGTTGTGTGAATGCAATGCAAGTCTTTT 60
QY 1738 GCCAAGTTCAGCTTATTTGACAGATGTGATGAACACCTCAATTTTCCCAAC-GT 1796
DB 61 GCCAAGTTCAGCTTATTTGACAGATGTGATGAACACCTCAATTTTCCCAACNGT 120
QY 1797 ATTCCAGC-GAAGTCAGTGAAGATGATGACCTAATGACCTAAAGTGGGCAATGTGACTG 1855
DB 121 ATTCCAGCNGAAGTCAGTGAAGATGATGACCTAATGACCTAAAGTGGGCAATGTGACTG 180
QY 1856 CCAAGATTCAGAAAGTCTGACATTAATGACCTGAGAGGAGACACAAAGAGTGGC 1915
DB 181 CCAAGATTCAGAAAGTCTGACATTAATGACCTGAGAGGAGACACAAAGAGTGGC 240
QY 1916 TTTAAATTCACAC--GTGACTGGTGAATCTTATGTTGGCTCCATTTGACAGAGAGC 1973
DB 241 TTTAAATTCACACCGGACGACGTCGATCTTTAGTGTGGCTCCATTTGACAGAGAGC 300
QY 1974 CGAAGTCATATCGGATCAAGTGTGGCACAAGAGTGGGAGGCTCTTTTAAAGCTC 2033

DB 301 CGAAGTCATATCGGATCAAGTGTGGCACAAGAGTGGGAGGCTCTTTGACGCTC 360
QY 2034 TGTGTGAGATTCACCTGATCTTATGATGATGAATGACACACCTCCAGAGTACGCA 2093
DB 361 TGTGTGAGATTCACCTGATCTTATGATGATGAATGACACACCTCCAGAGTACGCA 420
QY 2094 GGAATACAGGAGCTTGTCTTTCGCAATCCCTCAGTACGACCTGGAAGTCTCA-TTTTCG 2152
DB 421 GGAATACAGGAGCTTGTCTTTCGCAATCCCTCAGTACGACCTGGAAGTCTCAATTTTCG 480
QY 2153 AGGCTACTGATGATGATGACACTTATTTGGGGGTCCCAATTTTACATTT 2202
DB 481 AGGCTACTGATGATGATGACACTTATTTGGGGGAGGCCATTTTACATTT 530
RESULT 13
BM799332
LOCUS 472 bp mRNA linear EST 05-MAR-2002
DEFINITION K-EST0083174 S22SNUI6 Homo sapiens cDNA clone S22SNUI6-11-C06 5',
mRNA sequence.
ACCESSION BM799332
VERSION BM799332.1 GI:19147564
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 472)
AUTHORS Oh, K.J., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
CONTACT: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boseon-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribd.re.kr
Plate: 11 row: C column: 06
High quality sequence stop: 472.
Location/Qualifiers
1..472
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="S22SNUI6-11-C06"
/clone_lib="S22SNUI6"
/sex="F"
/tissue_type="ascites"
/cell_type="lymphoblast-like"
/cell_line="SNU-16"
/lab_host="DH10B"
/note="Organ: Stomach; Vector: pT73-Pac; Site 1: EcoRI;
Site 2: NotI. The S22SNUI6 library was contributed by the
Soares laboratory and it was constructed as described by
Bonaldo, M.F., Lennon, G. and Soares, M.B. (1996), Genome
Research 6(9): 791-806. RNA was prepared from harvested
cells of SNU-16 culture. SNU-16 cell was obtained from
Korean Cell Line Bank (KCLB). SNU-16 cell was established from
ascitic fluids of Korean patients by Park J.G. et al.
(1990), Cancer Res 50: 2773-2780."
BASE COUNT 102 a 119 c 90 g 161 t
ORIGIN
Query Match 14.0%; Score 467.2; DB 14; Length 472;
Best Local Similarity 99.4%; Pred. No. 9.1e-80;
Matches 469; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2667 GTGCAATTAATATTTTAAACAGATATTCCTCTTGTCTTAATATTTGCTAAATATT 2726
DB 1 GTGCAATTAATATTTTAAACAGATATTCCTCTTGTCTTAAATATTTGCTAAATATT 60

QY 2727 CTTTGTGAGTGGAGTCTGCTGTCGCGCAGGCTGAGTACAGTGTATCCAG 2786
 |||||
 DB 61 CTTTGTGAGTGGAGTCTGCTGTCGCGCAGGCTGAGTACAGTGTATCCAG 120
 |||||
 QY 2787 CTCACGTCAACCTCCGCTCTGCGGTTCACTGATTCCTGCTCAGCTTCTTAAG 2846
 |||||
 DB 121 CTCACGTCAACCTCCGCTCTGCGGTTCACTGATTCCTGCTCAGCTTCTTAAG 180
 |||||
 QY 2847 CTGGGTTTACAGCACCACCACCATGCGCACTTAATTTTATTAATAGACGG 2906
 |||||
 DB 181 CTGGGTTTACAGCACCACCACCATGCGCACTTAATTTTATTAATAGACGG 240
 |||||
 QY 2907 GGTTCGCAATTTGGCCAGGCTGCTGTAAGTCTGACGCAAGTATGCTGCTT 2966
 |||||
 DB 241 GGTTCGCAATTTGGCCAGGCTGCTGTAAGTCTGACGCAAGTATGCTGCTT 300
 |||||
 QY 2967 GGTTCGCAATTTAGCAGATGACCACTGACACCACTTAATTAATTAATGCTATA 3026
 |||||
 DB 301 GGTTCGCAATTTAGCAGATGACCACTGACACCACTTAATTAATTAATGCTATA 360
 |||||
 QY 3027 GACATTAGAGATTTTTCATTTTCCATGACATTTTCTCTGCAATGCTTACT 3086
 |||||
 DB 361 GACATTAGAGATTTTTCATTTTCCATGACATTTTCTCTGCAATGCTTACT 420
 |||||
 QY 3087 ACTGTGTTTTCCTTTTGGGCAAGACACTTAATTAATTAATGCTATA 3138
 |||||
 DB 421 ACTGTGTTTTCCTTTTGGGCAAGACACTTAATTAATTAATGCTATA 472
 |||||

RESULT 14

LOCUS BG981361 486 bp mRNA linear EST 12-JUN-2001
 DEFINITION CM4-CN0089-070201-746-g10 CN0089 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG981361
 VERSION BG981361.1 GI:14384096
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 486)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

COMMENT

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/br/cripte/gethtml2.pl?cl=CM4ct2-CM4-CN0089-
 070201-746-g10&c3=2001-02-07&c4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 3
 High quality sequence stop: 482.

FEATURES

1..486
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CN0089"
 /dev_stage="Adult"
 /note="Organ: colon_normal; Vector: puc18; Site_1: SmaI;

BASE COUNT

134 a 95 c 104 g 153 t

ORIGIN

Query Match 13.9%; Score 464.2; DB 13; Length 486;
 Best Local Similarity 99.4%; Pred. No. 3.4e-79;
 Matches 466; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

QY 1273 ATCCAACTGAGTGGAAATGTTACAGTTAGCTAAACAGTCTTGAAGAAGCAATACT 1332
 |||||
 DB 486 ATCCAACTGAGTGGAAATGTTACAGTTAGCTAAACAGTCTTGAAGAAGCAATACT 427
 |||||
 QY 1333 CCTCAGTACAACCTTAACGATAGAGTGTCTGACCAAGATTTCAAGACCTTTGTTG 1392
 |||||
 DB 426 CCTCAGTACAACCTTAACGATAGAGTGTCTGACCAAGATTTCAAGACCTTTGTTG 367
 |||||
 QY 1393 CAAATCAAGTTATTTGATATCAATGATCCATCTTTGAAATCAATTTATGA 1452
 |||||
 DB 366 CAAATCAAGTTATTTGATATCAATGATCCATCTTTGAAATCAATTTATGA 307
 |||||
 QY 1453 AACCTGACTTGTCTGAGACCAAAATTGGGTCACCATCTTAACATCCAGCCACT 1512
 |||||
 DB 306 AACCTGACTTGTCTGAGACCAAAATTGGGTCACCATCTTAACATCCAGCCACT 247
 |||||
 QY 1513 GATGCTGATGAGCCATTACTGAGAGTTCTAAATTTCTGATCATATCAATTAAGGAGAC 1572
 |||||
 DB 246 GATGCTGATGAGCCATTACTGAGAGTTCTAAATTTCTGATCATATCAATTAAGGAGAC 187
 |||||
 QY 1573 AGTGAGGAGACGCTGGGGGTTGACACAGATCCCATACCAACCGGATATGTCATATT 1632
 |||||
 DB 186 AGTGAGGAGACGCTGGGGGTTGACACAGATCCCATACCAACCGGATATGTCATATT 127
 |||||
 QY 1633 AAAAGGCTCTGATTTTGAACAGCAGCTGTTCCAACTTGTGTTCAAGAGAAAT 1692
 |||||
 DB 126 AAAAGGCTCTGATTTTGAACAGCAGCTGTTCCAACTTGTGTTCAAGAGAAAT 67
 |||||
 QY 1693 CTTGAGGCTCTGATTTTGAACAGCAGCTGTTCCAACTTGTGTTCAAGAGAAAT 1741
 |||||
 DB 66 CTTGAGGCTCTGATTTTGAACAGCAGCTGTTCCAACTTGTGTTCAAGAGAAAT 18
 |||||

RESULT 15

LOCUS BF334481 477 bp mRNA linear EST 22-NOV-2000
 DEFINITION M81-CN0355-160200-002-g05 CN0355 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF334481
 VERSION BF334481.1 GI:11305333
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 477)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE

20202663

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPSP/LICR Human Cancer Genome Project. This entry can be seen in the following URL:
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR1ct2-MR1-CT0355-160200-002-g05&t3=2000-02-16&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 41.

FEATURES

SOURCE

1..477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from 'ORESTES PCR' (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 160 a 104 c 95 g 118 t
ORIGIN

Query Match

13.5%; Score 452.8; DB 12; Length 477;

Best Local Similarity 97.5%; Pred. No. 5,3e-77; Matches 460; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1021 GCAAGAGTATGATGCGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 1080
DB 6 GCAATGATGAGGACGAAAAACCACTTTCATATCCGCTGGAATTCATGTAAAGTTAA 65
QY 1081 GATATTATATGATATCCACTTACATGTCCTGACCAATACCGTATTGAGTCCAGAG 1140
DB 66 AGATATATATATATATCCACTTACATGTCCTGACCAATACCGTATTGAGTCCAGAG 125
QY 1141 AATGAACGACTGGGTACAGATGCGGACCCTTACTGCAATGACAGGATGAAGAAAT 1200
DB 126 AATGAACGACTGGGTACAGATGCGGACCCTTACTGCAATGACAGGATGAAGAAAT 185
QY 1201 ACTGCCAAGATTTTCTTAACTACAGATTTGTGAGCAAACTCCAAACTTCCATGAT 1260
DB 186 ACTGCCAAGATTTTCTTAACTACAGATTTGTGAGCAAACTCCAAACTTCCATGAT 245
QY 1261 GGACTCTTCTTATCCAACTTATGCTGGAATGTTACGTTAAGTAAACGTCCTTGAAG 1320
DB 246 GGACTCTTCTTATCCAACTTATGCTGGAATGTTACGTTAAGTAAACGTCCTTGAAG 305
QY 1321 AAGCAAGATACCTCTCAGTACAACTTAAAGTAGAGGTCTGACAAAGATTCAAGACC 1380
DB 306 AAGCAAGATACCTCTCAGTACAACTTAAAGTAGAGGTCTGACAAAGATTCAAGACC 365
QY 1381 CTTTGTGTTTGCAAAATCAAGTTATGATATCAATGATCAATCCCATCTTTGAAAAA 1440
DB 366 CTTTGTGTTTGCAAAATCAAGTTATGATATCAATGATCAATCCCATCTTTGAAAAA 425
QY 1441 TCAGATTATGAAACCTGACTCTTGTGTAAGACACAAACATTGGGTCCACCA 1492
DB 426 TCAGATTATGAAACCTGACTCTTGTGTAAGACACACAAATTGGGTCCACCA 477

Search completed: June 22, 2003, 00:52:25
Job time : 4134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compygen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:08:50 ; Search time 2438 Seconds

(without alignments)
5526.926 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIQAHLSHLCULMLYATG.....DKGXNVSQAQSEVPLRS 832

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlj
-Q=/cgr2.1/USPTO/spool/US10025380/runat_17062003_173328_24678/app_query.fasta_1.1031
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=dot -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10025380 @CCN_1_1.2087 @runat_17062003_173328_24678 -NCPU=6 -ICPU=3
-NO MAP -LARGEDEV= -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hci:*
9: gb_est1:*
10: gb_est2:*
11: gb_hci:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1227	28.4	1125	14	BM918123	BM918123 AGENCOURT
2	1046	24.2	613	10	AM859603	AM859603 MR1-CT035
3	1032	23.9	625	12	BE696136	BE696136 MR1-CT025
4	955	22.1	573	9	A1347909	A1347909 gp60e04.x
5	884	20.5	873	12	BF582721	BF582721 602094137
6	881	20.4	540	13	BG978653	BG978653 PM1-CN009
7	863	20.0	514	9	A1733765	A1733765 z190f07.y
8	856	19.8	590	10	AM859530	AM859530 MR1-CT035
9	853	19.7	519	10	AM859550	AM859550 MR1-CT035
10	839	19.4	530	9	AA102326	AA102326 z190f07.y
11	819	19.0	657	12	BF584025	BF584025 602096265
12	791	18.3	486	13	BG981361	BG981361 MR1-CT024
13	781	18.1	495	10	AM853021	AM853021 MR1-CT035
14	779	18.0	477	12	BF334481	BF334481 MR1-CT035
15	777	18.0	563	10	AM862312	AM862312 RC4-CT036
16	772	17.9	753	12	BF577684	BF577684 602092567
17	763.5	17.7	620	10	AM377543	AM377543 OV0-CT022
18	750	17.4	455	12	BE696126	BE696126 MR1-CT025
19	734	17.0	507	10	AM753130	AM753130 RC1-CT024
20	726	16.8	551	12	BE808624	BE808624 213876 MA
21	719	16.6	464	10	AM391746	AM391746 CM2-ST022
22	702	16.2	406	10	AM351846	AM351846 RC1-CT019
23	702	16.2	508	9	AA871508	AA871508 vq36c11.x
24	674	15.6	436	13	BG981357	BG981357 CM4-CN008
25	666	15.4	461	10	AM859489	AM859489 MR1-CT035
26	663	15.3	403	9	AA053188	AA053188 z172a06.x
27	660.5	15.3	417	10	AM351659	AM351659 QV2-CT014
28	652	15.1	405	13	BG980846	BG980846 CM4-CN008
29	645	14.9	486	9	A1641856	A1641856 vq36c11.y
30	644	14.9	408	12	BF754330	BF754330 IL5-CT051
31	639	14.8	423	9	AA308216	AA308216 EST179051
32	636	14.7	435	13	BG981295	BG981295 CM4-CN008
33	635	14.7	447	10	AM859567	AM859567 MR1-CT035
34	615	14.3	395	10	AM578054	AM578054 RC3-CN001
35	615	14.2	477	10	AM859516	AM859516 MR1-CT035
36	600	13.9	431	13	BG981285	BG981285 CM4-CN008
37	594	13.7	373	13	BG980900	BG980900 CM4-CN008
38	591	13.7	383	13	BG980835	BG980835 CM4-CN008
39	587	13.6	370	12	BF754194	BF754194 IL5-CT051
40	586	13.6	377	10	AM753432	AM753432 MR1-CT025
41	583	13.5	410	13	BG982186	BG982186 CM4-CN008
42	570	13.2	367	13	BG982184	BG982184 CM4-CN008
43	544	12.6	419	10	AM375303	AM375303 QV2-CT014
44	536	12.4	347	10	AM351613	AM351613 QV2-CT014
45	528	12.2	321	13	BG980647	BG980647 CM4-CN008

ALIGNMENTS

RESULT 1
LOCUS BM918123 1125 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT 6611470 NIH_MGC_106 Homo sapiens CDNA IMAGE:5485497
5' mRNA sequence.
ACCESSION BM918123
VERSION BM918123.1 GI:19368502
KEYWORDS
SOURCE
ORGANISM Homo sapiens
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsb@nih.gov
Tissue Procurement: Dr. Daniel McVicar, DBS/NCI
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2016 row: h column: 10
High quality sequence stop: 662.
Location/Qualifiers

FEATURES

1. 1125
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5485497"
/clone_lib="NIH MGC 106"
/issue_type="natural killer cells, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: Blood; Vector: pOT57; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC library."

BASE COUNT 259 a 321 c 258 g 286 t 1 others
ORIGIN

Alignment Scores:

Prod. No.: 2.87e-131 Length: 1125
Score: 1227.00 Matches: 240
Percent Similarity: 96.41% Conservative: 2
Best Local Similarity: 95.62% Mismatches: 8
Query Match: 28.40% Indels: 1
DB: 14 Gaps: 0

US-10-025-380-1081 (1-832) x BW918123 (1-1125)

QY 580 AlaileGlyThrlyValGlyAsnValThrAlaIlyAspProGluGlyLeuAspIleSer 599
DB 1 GCTATGCGCTAAAGGGGCAATGATCTGCCAAGATCCAGAAGTCTGGACATTAAGC 60
QY 600 TySerleuArgGlyAspThrArgGlyTyrLeuIlyleAspHisValThrGlyGluIle 619
DB 61 TATTCACGTGAGGGAGACACAGAGGTGGCTTAAATTGACACAGCTGGTGAATC 120
QY 620 PheSerValAlaProIleuAspArgGlyAlaGlySerProTyrArgValGluValAla 639
DB 121 TTATAGTGGCTCATTTGACAGAGAACCGAGATCATATCGGGACAGAGTGAGG 180
QY 640 ThrGluValGlyIlySerSerleuSerValSerGluPheHisleuIleuMetAsp 659
DB 181 ACAGAGTACGGGGGCTTCTTAAAGTCTGTGCAAGTTCACCTGATCTTAAGAT 240
QY 660 ValAsnAspAsnProProArgLeuAlaIlyAspTyrThrGlyLeuPheCysHisPro 679
DB 241 GTGAATGACAAACCTCCACAGGTACCAAGACATACAGGGCTTCTTCGCAATCCC 300
QY 680 LeuSerAlaProGlySerleuIlePheGluAlaThrAspAspArgHisleuPheArg 699
DB 301 CTAGTGCACCTGAAAGTCTCATTTTGAAGCTACGTATATATGACACTTAATTTCCG 360
QY 700 GlyProHisPheThrPheSerleuGlySerGlySerleuGluAsnAspTyrGluValSer 719
DB 361 GGTCCCATTTTACATTTTCCCTCGGACGAGGAAGCTTACAAACGACTGGGAAGTTCC 420
QY 720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluValArg 739
DB 421 AAATCAATGTAATCATCTATCCGACGTGTACACGAGCACAAGACTTTGAGAGAGGGC 480

QY 740 TyrValIleuIleAlaGlyLeuAsnAspGlyIlyArgProProLeuGluGlyIleValSer 759
DB 481 TATGTGCTTATATCGCATCAATATGGGGGTCCGCCACCTTGGAAAGCATTTGTTCT 540
QY 760 LeuProValThrPheCysSerCysValGlyIlySerCysPheArgProAlaGlyHisGln 779
DB 541 TTACAGTATACCTTCGACGTTGTGTGGAGAAAGTGTTCCTCCGACAGTACCGAG 600
QY 780 ThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIle 799
DB 601 ACTGGGATACCACTGTGGGCAAGGCGAGTTGATCTCTACACACCTTCTGGATG 660
QY 800 GlyIleIleLeuAlaValAlaPheIleArgIleIlyIlyAspTyrGlyIlyAspAsnVal 819
DB 661 GGCATATTTTATCCGACGTTGGCGGCTATCCGCTTAAAGAGACTACCGCCAGATATGT 720
QY 819 IgluSerAlaGluAlaIlySerGluValIlyPro 829
DB 721 TGACGCTCCTACGCTCTTGAACGCCAACCC 751

RESULT 2
LOCUS AM859603 613 bp mRNA linear EST 19-MAY-2000
DEFINITION MR1-CT0355-180200-007-e07 CT0355 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM859603
VERSION AM859603.1 GI:7955296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 613)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=MR1-CT0355-180
200-007-e07&ct3=2000-02-18&ct4=1)
Seq primer: puc 18 forward
High quality sequence stop: 613.
Location/Qualifiers

FEATURES

1. 613
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0355"
/dev_stage="Adult"
/note="Organ: Colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 195 a 125 c 134 g 159 t
ORIGIN
Alignment Scores:

Pred. No.: 1.13e-110 Length: 613
 Score: 1046.00 Matches: 202
 Percent Similarity: 100.0% Conservative: 2
 Best Local Similarity: 99.02% Mismatches: 0
 Query Match: 24.21% Indels: 0
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AW859603 (1-613)

QY 411 LysleGlnAspThrProGlnIYrAsnLeuThrIleGluValSerAspLysAspPheLys 430
 Db 1 AAGAGGAGAGTACTCTCTGAGTACTTAAACGATAGAGGTGTCTGCAAGATTTCAG 60
 QY 431 ThrLeuCyaspheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlu 450
 Db 61 ACCCTGTTTGTGGCAATCAACGATTATGATCATATCATGATCCCATCTTGAA 120
 QY 451 LysSerAspIYrGluValLeuThrLeuValIleGluAspThrAsnIleGlySerThrIleLeu 470
 Db 121 AATACAGATTATGAAACCTGACCTTCTGAAAGACCAAACTTGGGTCCACCATCTTA 180
 QY 471 ThrIleGlnAlaThrAspAlaSerGluProPheThrGlySerSerIYrIleLeuYrHis 490
 Db 181 ACCATCCAGCCCACTGATGCTGATGAGCCATTTACTGGAGTCTTAATTCGATCAT 240
 QY 491 IleIleIYrGluYrAspSerGluGlyArgLeuGlyValAspThrAspProIleIsthrAsnThr 510
 Db 241 ATCATAAAGGAGACGATGAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACACC 300
 QY 511 GlyIYrValIleIleIYrValProLeuAspPheGluThrAlaValSerAsnIleVal 530
 Db 301 GGATATGTCTAATTAATTAAGCCCTTGATTTGAAACACAGCTGTTTCAACATTTGTG 360
 QY 531 PheIYrAlaGluAsnProGluProLeuValPheGlyValIYrYrAsnAlaSerSerPhe 550
 Db 361 TTCAAGCAAGAAATCTGAGCTTCTGATGTTGGTGTGAAGTCAATGCAAGTCTTTT 420
 QY 551 AlaIYrPheThrLeuIleValIsthrAspValAsnGluValProGlnPheSerGlnHisVal 570
 Db 421 GCCAGTTTCAGCTTATGTCAGCAGATGATGATGACACTCATTTTCCCAACAGTA 480
 QY 571 PheGlnAlaIYrValSerGluAspValAlaIleGlyThrIYrValGluValIsthrAla 590
 Db 481 TTCCAGCGCAAGATCTGAGTGTAGCTATAGGCACTAAGTGGCAATGAGACTGCC 540
 QY 591 LysAspProGluGlyLeuAspIleSerIYrSerIleArgIYrAspThrArgGlyYrProLeu 610
 Db 541 AAGGATCCAGAAAGTCTGACATTAAGCTATTCACTGAGGAGACACAAAGAGGTGGTT 600
 QY 611 LysIleAspHis 614
 Db 601 AAAATTGAAACAC 612

RESULT 3
 BE696136/c 625 bp mRNA linear EST 11-SEP-2000
 LOCUS BE696136 MRI-CT0258-040700-002-e10 CT0258 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE696136
 VERSION BE696136.1 GI:10083296
 ACCESSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 625)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?fl=ec2=MRI-CT0258-040
 700-002-e10&ct=2000-07-04&ct=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 620.

FEATURES

source
 1. 625
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="CT0258"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived
 from ORSTS PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the puc 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."
 BASE COUNT 143 a 141 c 153 g 188 t
 ORIGIN

Alignment Scores:

Pred. No.: 4.94e-109 Length: 625
 Score: 1032.00 Matches: 202
 Percent Similarity: 98.54% Conservative: 1
 Best Local Similarity: 98.06% Mismatches: 3
 Query Match: 23.88% Indels: 1
 DB: 12 Gaps: 0

US-10-025-380-1081 (1-832) x BE696136 (1-625)

QY 60 PheGluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGluLeuYr 79
 Db 617 TTGGAAGTAAAC-GGGAGACAGACACATATTGTGATGAAAGGAGGACTCTGTAT 559
 QY 80 TyrAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaLeu 99
 Db 558 TACAAAGAGCCCTTGACAGGAGCAACAGATCTACCAATCTCCAGGTTGACGCCCTG 499
 QY 100 AspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIsthrAla 119
 Db 498 GACGCTAATGAAATTAATGATGAGGAGTCCAGTCCCTATACCAATGAAGAGGACATC 439
 QY 120 AsnAspAsnArgProThrPheLeuGlnSerIYrGlyGlySerValArgGlnAsnSer 139
 Db 438 AACGAAATCGAACCCAGCTTCTCCAGTCAAAAGTCAAGAGCTCAAGTAAAGCAACTCT 379
 QY 140 ArgProGlyIYrAspPheLeuYrValAsnAlaThrAspLeuAspAspProAlaThrPro 159
 Db 378 CGCCCAAGAAAGCCCTTCTTGTATGATGACACAGACTGGATATCCGGCACTCC 319
 QY 160 AsnGluGlnLeuYrIYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetYr 179
 Db 318 AATGGCCAGCTTATTAACAGATGTGATCCAGCTTCCATATCAACAATGTTATGTAC 259
 QY 180 PheGlnIleAsnAsnIYrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeu 199
 Db 258 TTTCAGATCAACAACAACGAGGAGCACTCTCTTAACCCAGAGGAGATCTCAGAAATG 199
 QY 200 AsnProAlaIYrAsnProSerIYrAsnLeuValIleSerValIYrAspMetGlyGln 219
 Db 198 AATCGCTAAGAAATCTCTTATATCTGTGATCTCAGTGAAGACATGGAGGCCAG 139

QY	220	SequiuaensepseaseApThThreValaspielleValThnGlaasnilleTP	239
Db	138	AGTGAAGATTCCTTCAAGTATCCACATCTGTGATATCAAGTACAGAGAAATTTTG	79
QY	240	LysalaprolyspProValgluMetValgluanserThrapProHlePiolleyle	259
Db	78	AAAGACCAATVACCTGTGGAGATGGTGGAAACTCACTGATCTCTCAACCCATCCAAATC	19
QY	260	ThrglnValargTPan	265
Db	18	ACTCAAGTCGGTGGAAAT	1
RESULT 4			
LOCUS	AI347909/c	573 bp	mRNA linear EST 30-DEC-1998
DEFINITION	gpc0e04.x1 NCI CGAP C08 Homo sapiens CDNA clone IMAGE:1927422.3'		
ACCESSION	AI347909		similar to Y15336 Q15336 LI-CADHERIN. [1] ;, mRNA sequence.
VERSION	AI347909		
KEYWORDS	AI347909.1 GI:4085115		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	1 (bases 1 to 573)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-rt@mail.nih.gov		
	Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.		
	Emmert-Buck, M.D., Ph.D.		
	CDNA Library Preparation: M. Bento Soares, Ph.D.		
	CDNA Sequencing Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/ILNI at:		
	www.bio.lnlnl.gov/bbrp/image/image.html		
	Seq primer: -40UP from Gibco		
	High quality sequence Etop: 483.		
FEATURES	Location/Qualifiers		
source	1..573		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1927422"		
	/clone_11b="NCI CGAP_C08"		
	/tissue_type="adenocarcinoma"		
	/lab_host="DH10B"		
	/note="Organ: colon; Vector: pUT73D-Pac (Pharmacia) with a		
	modified polylinker; 1st strand cDNA was prepared from		
	colon adenocarcinoma, and was then primed with a Not I -		
	oligo (dt) primer. Double-stranded cDNA was ligated to Eco		
	RI adaptors (Pharmacia), digested with Not I and cloned		
	into the Not I and Eco RI sites of the modified pUT73		
	vector. Library is normalized. Library was constructed by		
	Bento Soares and M. Felina Bonaldo. "		
BASE COUNT	151 a 155 c 139 g 128 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	3, 82e-100	Length:	573
Score:	955.00	Matches:	183
Percent Similarity:	98.94%	Conservative:	3
Best Local Similarity:	97.34%	Mismatches:	1
Query Match:	22.10%	Indels:	1
DB:	9	Gaps:	0
US-10-025-380-1081 (1-832) x AI347909 (1-573)			
QY	613	ApPhHValaThThGlyGluIlePheSerValAlaProLeuApArgGluAlaGlySerPro	632

Db	Accession	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	572	GAACACGTGACTGTGAGATCTTTAGTGTGGCTCCATTTGACAGAGAAAGCCGGAATGCCA	513					
Qy	633	TYATyValGlnValValAlaThrGlnValGlyGlySerSerSerSerValSerGlu	652					
Db	512	TATCGGGTACAAAGTGTGGCCACAGAAAGTGGGGGGCTTCTTAACTGTGTGCAG	453					
Qy	653	PhnHisLeuLeuLeuMetAspValAlaAspAsnPro-ProArgLeuAlaIysAspTyrThr	672					
Db	452	TTCCACTGTATCCTTAGATGATGATGATGACAACTTCCAGAGTTAGCCAGACATCAC	393					
Qy	672	IGLYLeuPhePheCysHisProLeuSerAlaProGlySerLeuIlePheGlnAlaThrAs	692					
Db	392	GGGCTGTCTTCTTGCCATCCCTCCAGTGCACCGAAGCTCATTTTCAGAGCTACTGA	333					
Qy	692	PAAPAspGlnHisLeuPheArgGlyProHisPheThrPheSerLeuGlySerGlySerIle	712					
Db	332	TGATGATCAGACACTTATTTCCGGGGGTCCCAATTTTACATTTTCCCTCGCAGTGGAACTT	273					
Qy	712	UGAlaAspAspTrrGlnValSerIysIleAsnGlyThrHisAlaArgLeuSerThrArgHis	732					
Db	272	ACAAAGAGACTGGAGAGATTTCACAAATCAATGGTACTCAGTCCGAGCTGTCTACCAAGCA	213					
Qy	732	SThrAspPheGlnGluGluArgAlaTyrValValLeuIleArgIleAsnAspGlyGlyArgPr	752					
Db	212	CACAGACTTGGAGAGAGAGGGCGCTATGTCGTCTTGATCCGATCAATATGAGGGGTCCGCC	153					
Qy	752	OProLeuGlnGlyIleValSerLeuProValThrPheCysSerCysValGluGlySerC	772					
Db	152	ACCTTGAGAGAGCACTTGTCTTTACAGATTACATTCGTGCAGTGTGTGGAGAGAACTTG	93					
Qy	772	SPheArgProAlaGlyHisGlnThrGlyIleProThrValGlyMetAspValaGlyIleLe	792					
Db	92	TTTCCGGCCAGAGGTACCAAGACTGGGATPCCACATGTGGGCAATGGCAGTTGGTATACT	33					
Qy	792	ULeuThrThrLeuLeuValIle 799						
Db	32	GCTGACCAACCATTTCTGATGTTG 11						
RESULT 5	BF582721	873 bp	mRNA	linear	EST 12-DEC-2000			
LOCUS	602094137F1	NCI_CGAP_Co24	Mus musculus	CDNA clone	IMAGE:4208360 5',			
DEFINITION	mRNA sequence.							
ACCESSION	BF582721							
VERSION	BF582721.1	GI:11656439						
KEYWORDS	EST.							
SOURCE	house mouse.							
ORGANISM	Mus musculus.							
REFERENCE	Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;							
AUTHORS	Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.							
TITLE	1 (bases 1 to 873)							
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.							
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)							
	Unpublished (1999)							
	Contact: Robert Strausberg, Ph.D.							
	Email: cgabs-rc@mail.nih.gov							
	Tissue Procurement: Jeffrey E. Green, M.D.							
	CDNA Library Preparation: Life Technologies, Inc.							
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)							
	DNA Sequencing by: Incyte Genomics, Inc.							
	Clone distribution: MGC clone distribution information can be							
	found through the I.M.A.G.E. Consortium/LNLN at:							
	http://image.lnl.gov							
	Plate: LLM9772 row: j column: 09							
	High quality sequence stop: 652.							
FEATURES	Location/Qualifiers							
source	1..873							
	/organism="Mus musculus"							
	/strain="FVB/N"							
	/db_xref="taxon:10090"							
	/clone="IMAGE:4208360"							
	/clone_lib="NCI CGAP Co24"							
	/lab_host="DH10B (TI phase-resistant)"							

Db 418 GACATAGCTATTCTAGAGGGAGACACAGAGTTGGCTTAATATGACACGTGACT 359
QY 617 GYGLuilePheSerValAlaProLeuAspArgGluAglYserProTYraGValGln 636
Db 358 GGTGAGATCTTTAGTGTGGCTCCATTGACAGAGAACCCGGAAGTCCATATCGGGTACAA 299
QY 637 ValValAlaThrGluValAGlyYserSerLeuSerSerValSerGluPheHileuile 656
Db 298 GTGTGGCCACAGAGAGAGGGGGGTCTTCTTAAGCTCTGTGTCAGAGTTCCACCTGATC 239
QY 657 LeuMetAspValAlaAspAspProProArgLeuAlaValAspTYrThrGlyLeuPhePhe 676
Db 238 CTTATGATGATGATGATGACACCTCCAGGCTACCAAGACATACACAGGCTTGTCTTC 179
QY 677 CyHlePheLeuSerAlaProGlySerLeuilePheGluAlaThrAspAspArgGlnHle 636
Db 178 TGCATCCCTCAGTCACCTGACCTGAGTCAATTTTCAGGCTACTGATGATGATCAGCAC 119
QY 697 LeuPheArgGlyProHlePheThrPheSerLeuGlyYserGlySerLeuGlnAspAspTYr 716
Db 118 TTATTTGGGGTCCCTTTTACATTTTCCCTCGGCGAGTGAAGCTTACAAAGCACTGG 59
QY 717 GluValSerLYsIleAenglyThrHleAlaArgLeuSerThrArgHleThrAsp 734
Db 58 GAAGTTTCCAATC-AATGGTACTCATGCCGACTGTCCACATGACATGAT 6
RESULT 7
LOCUS A1733765 514 bp mRNA linear EST 14-JUN-1999
DEFINITION z190f07.y5 Stratagene colon (#937204) Homo sapiens CDNA clone
IMAGE:511909 5' similar to TR:Q12864 Q12864 INTESTINAL
PEPTIDE-ASSOCIATED TRANSPORTER HPT-1. [1] ;, mRNA sequence.
ACCESSION A1733765
VERSION A1733765.1 GI:5054878
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 514)
AUTHORS NCI/NIH-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: z190f07.x5
Contact: Robert Strauberg, Ph.D.
Email: cgapbs-i@mail.nih.gov
This clone is available royalty-free through LANT; contact the
IMAGE Consortium (infoimage.lnl.gov) for further information.
Original clone citation: Washu-Merck EST Project
This read has been verified (found to hit its original self in the
correct orientation)
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gdbco
High quality sequence stop: 405.
FEATURES
source
Location/Qualifiers
1. 514
/organism="Homo sapiens"
/db_xref="GDB:384404"
/db_xref="taxon:9606"
/clone="IMAGE:511909"
/clone_lib="Stratagene colon (#937204)"
/issue_type="tumor"
/tissue_type="Colon"
/cell_line="T84 carcinoma cell line"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: colon; Vector: pBluescript SK-; Site 1:
EcORI, Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor
sequence: 5' GAATTCGACGACGAG 3' ~3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3'."

BASE COUNT 135 a 115 c 132 g 133 t

ORIGIN
Alignment Scores:
Pred. No.: 1.57e-89 Length: 514
Score: 863.00 Matches: 168
Percent Similarity: 98.254 Conservative: 0
Best Local Similarity: 98.254 Mismatches: 3
Query Match: 19.974 Indels: 0
DB: 9 Gaps: 0
US-10-025-380-1081 (1-832) x A1733765 (1-514)
QY 531 PheLYsAlaGluAspProGluProLeuValPheGlyValLYsAsnAlaSerPhe 550
Db 2 TTCAAAGCAGAAATCTGAGCTTACGCTTACGTTGGTGAATACATGCAAGTCTTTT 61
QY 551 AlaLYsPheThrLeuileValThrAspValAenglyAlaProGlnPheSerGlnHleVal 570
Db 62 GCCAAGTTCACGCTTATTTGACAGATGGAAGACACCTCAATTTCCCAACACGTA 121
QY 571 PheGlnAlaLYsValSerGluAspValAlaileGlyThrLYsValGlyAsnValThrAla 590
Db 122 TTCCAAGCAGAAAGTCAAGTGAAGATGATGCTATGACCTAAAGTGGCAATGTGAGTGC 181
QY 591 LYsAspProGluGlyLYsAspIleSerTYrSerLeuArgLYsAspThrArgGlyTYrLeu 610
Db 182 AAGATCCAGAAAGCTGTGACATTAAGCTTTCATGAGGGAGACACAAAGAGTTGGCTT 241
QY 611 LYsIleAspHleValThrGlyGluilePheSerValAlaProLeuAspArgGluAglY 630
Db 242 AAAATGACACAGTGAAGTGAAGATCTTAATGTGGCTCCATTGACAGAGACCGCA 301
QY 631 SerProTYrArgValGlnValValAlaThrGluValAGlyYserSerLeuSerSerVal 650
Db 302 AGTCCATATCGGGTAAAGTGTGGCCACAGAGTGGGGGTCTTCTTGAGCTTGtg 361
QY 651 SerGluPheHleuileLeuMetAspValAlaAspAspProProArgLeuAlaLYsAsp 670
Db 362 TCAGAGTTCCACCTGATCTTATGATGATGATGACACCTCCAGGCTACCAAGAC 421
QY 671 TYrThrGlyLeuPhePheCyHlePheLeuSerAlaProGlySerLeuilePheGluAla 690
Db 422 TACACGGGCTTGTCTTCTTGCATCCCTCAGTGAACCTCAATTTCCAGGCT 481
QY 691 ThrAspAspAspGlnHleuPheArgGlyPro 701
Db 482 ACTGATGATGATCAGACATATTTTCGATCCA 514
RESULT 8
LOCUS AW859530 590 bp mRNA linear EST 19-MAY-2000
DEFINITION WRI-CT0355-160200-003-e05 CT0355 Homo sapiens CDNA, mRNA sequence.
ACCESSION AW859530
VERSION AW859530.1 GI:7955223
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 590)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,
Nagel,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Dongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR1-CT0355-160
200-004-b02kt3=2000-02-16&t=1)
Seq primer: puc 18 forward
High quality sequence start: 22
High quality sequence stop: 577.
Location/Qualifiers

FEATURES

source

1. 590
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 183 a 127 c 132 g 148 t

ORIGIN

Alignment Scores:

Pred. No.: 1,3e-88 Length: 590
Score: 856.00 Matches: 179
Percent Similarity: 97.30% Conservative: 1
Best Local Similarity: 96.76% Mismatches: 5
Query Match: 19.81% Indels: 3
DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AW859550 (1-590)

415 ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThrLeuGlySph 434
Db 19 ACCCCCTGATCACTGACGACGATGAGGTCTCTGACAAAGA-TTCAAGACCCCTTTT 77
Qy 435 ValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSerAspTyr 454
Db 78 GTGCAATCAACGCTATGATATCATGATCATGATCCCATCTTGAATAATCATGATTA 137
Qy 455 GYAAsnLeuThrLeuIleGluAspThrAsnIleGlySerThrIleLeuThrIleGlnIle 474
Db 138 GGAACCTGACTCTTCTGTAAGACACAAACATGGGTCCACCATCTTAACCATCCAGGCC 197
Qy 475 ThrAspAlaAspGluProPheThrGlySerSerLysIleLeuThrIleIleIleGly 494
Db 198 ACTGATGCTGATGACCATTTACTGGAGTTCTTAATAATCTGTATCATATCATTAAGGGA 257
Qy 495 AAspSerGluGlyArgLeuGlyValAspThrAspProIleThrAsnThrGlyTyrValIle 514
Db 258 GACAGTGAAGGAGCGCTGGGGGTTGACACAGATCCCATACCAACACCGGATATGTCAT 317
Qy 515 IleValGlyProLeuAspPheGluThrIleAlaValSerAsnIleValPheLysAlaGlu 534
Db 318 ATTAATAAGCCCTCTGATTTTGAACAGACGCTGTTTCCAACTGTTGTTCAAGCAGAA 377
Qy 535 AAspProGluProLeuValPheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThr 554
Db 378 ATCTGAGCCCTTACTGCTGCTGGTGTGAAGTACAAAGTCTTTTCCCAAGTTCAGG 437
Qy 555 LeuIleValThrAspValAsnGluAlaProGlnPheSerGlnIleValPheGlnAlaLys 574
Db 438 CTTATGTGACAGATGTGATGATGACACCTCAATTTTCCCAACAGTATTCAGAGGAA 497
Qy 575 ValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLysAspProGlu 594
Db 498 GTCACTGAGATGTACTATAGCACTAAAGTGGGCCATGTGATGCCAAGGT-CCAGAA 556

Qy 595 GYLeuAspIleSer 599
Db 557 GGCTG-GACATAAGC 570

RESULT 9

AW859550 519 bp mRNA linear EST 19-MAY-2000
MR1-CT0355-160200-004-b02 CT0355 Homo sapiens CDNA, mRNA sequence.
AW859550
VERSION AM859550.1 GI:7955243
KEYWORDS EST.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 519)
Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

MEDLINE

COMMENT

Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR1-CT0355-160
200-004-b02kt3=2000-02-16&t=1)
Seq primer: puc 18 forward
High quality sequence start: 81
High quality sequence stop: 517.
Location/Qualifiers

FEATURES

source

1. 519
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1ib="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 160 a 111 c 107 g 141 t

ORIGIN

Alignment Scores:

Pred. No.: 2,33e-88 Length: 519
Score: 853.00 Matches: 166
Percent Similarity: 98.26% Conservative: 3
Best Local Similarity: 96.51% Mismatches: 3
Query Match: 19.74% Indels: 0
DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AW859550 (1-519)

415 ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThrLeuGlySph 434
Db 3 ACTCTGATTAACCAACGATTAAGGTGTCTGACAGATTAACAAACCCCTTTGTTT 62
Qy 435 ValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSerAspTyr 454

QY 416 ProGlnTyrAsnLeuThrIleGluValSerAspIleAspPheLeuThrLeuGlyCysPheVal 435
 DB 426 CCTCAGACCACTTAAAGATAGAGGTCTGACCAAGATTTCAGACCCCTTTGTTG 367
 QY 436 GlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnIleGlySerAspTyrGly 455
 DB 366 CAATCAACGTTATGTATCATATGATCATGATCCCATCTTTGAAAAATCAGATTATAGA 307
 QY 456 AsnLeuThrLeuValIleGluAspThrAsnIleGlySerThrIleLeuThrIleGlnIleValThr 475
 DB 306 AACCTGACTCTTCTGTAAGACACAAACATGGGTGCAACATCTTAACCATCCAGGCCACT 247
 QY 476 AspAlaAspGluProPheThrGlySerSerIleIleLeuThrIleIleIleGlyAsp 495
 DB 246 GATCTCATAGACCATTTACTGAGAGTTCTAAATTTCTGATCATATCAAAAGGAGAC 187
 QY 496 SerGlnIleValIleGluGlyValIleAspThrAspProIleThrAsnThrGlyTyrValIleIle 515
 DB 186 AGTAGAGGACGCTGGGGGTGACACAGATCCCATCCAAACCGGATATGTCATATT 127
 QY 516 LysIleProIleAspPheGlnIleThrAlaIleValSerAsnIleValPheIleValIleGluAsn 535
 DB 126 AAAAAGCCTCTTATTTTGAACAGCAGCTGTTCCAACTATGTTCAAGAGCAAAAT 67
 QY 536 ProGluProLeuValPheGlyValIleValIleValIleValIleValIleValIleValIleVal 550
 DB 66 CCTGAGCCTTACTGTTTGTGTGATGATCAATGCAAGTCTTTT 22
 RESULT 13
 AM853021/c 495 bp mRNA linear EST 19-MAY-2000
 LOCUS R01-CT0249-170200-025-b11 CT0249 Homo sapiens cDNA, mRNA sequence.
 DEFINITION AM853021
 ACCESSION AM853021.1 GI:7948538
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 495)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the PAPER/P/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1&t2=RC1-CT0249-170
 200-025-b11&t3=2000-02-17&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 495.
 FEATURES
 source location/Qualifiers
 1. 495
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CT0249"
 /dev_stage="Adult"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
 SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196
 716 - Ludwig Institute for Cancer Research) profiles
 into the pUC 18 vector. Reverse transcription of tissue
 mRNA and cDNA amplification were performed under low
 stringency conditions."

BASE COUNT 121 a 108 c 124 g 142 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5.04e-80 Length: 495
 Score: 781.00 Matches: 153
 Percent Similarity: 96.23% Conservative: 0
 Best Local Similarity: 96.23% Mismatches: 6
 Query Match: 18.07% Indels: 1
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x AM853021 (1-495)

QY 6 HisLeuHisSerLeuGlyLeuLeuMetLeuTyrLeuAlaThrGlyTyrGlyGlnGlnGly 25
 DB 493 CATCTTCACTCCCTGCTGCTCTTATGCTTATTTGGCACTGATATGCGAAGAGGGG 434
 QY 26 LysPheSerGlyProLeuIleProMetThrPheSerIleTyrGlnGlyGlnProSer 45
 DB 433 AATTTTAGTGAGCCCTGAAACCATGACATTTCTATTATTAAGGCCAAGACCGAGT 374
 QY 46 GlnIleIlePheGlnPheIleValIleGluArgGlnGlyLeuLeuTyrTyrAsnArgAlaLeuAsp 65
 DB 373 CAATTAATATTCAGTTAAGCCCAATCTCTGCTGATGCTTGAATTAACCTGGGAG 314
 QY 66 ThrAspAsnIlePheValIleGluArgGlnGlyLeuLeuTyrTyrAsnArgAlaLeuAsp 85
 DB 313 ACAGACAACTATTTGTGTATGATAGAGGAGGAGACTTCTGTATTAACAGACCTTGGAC 254
 QY 86 ArgGluThrArgSerThrHisAsnLeuGlnValAlaIleAspAlaAsnGlyIleIle 105
 DB 253 AGGAAACAAAGATCTACTACATCTCCAGTTGGACGCTTGGAGCTTAATGAAATTATA 194
 QY 106 ValGlnGlyProValProIleThrIleGluValIleAspIleAsnAspAsnArgProThr 125
 DB 193 GTGAGAGGTCCAGTCCCTTATCCATAGAGGAGGAGACATCAACACATGACACCCAG 134
 QY 126 PheLeuGlnSerLeuTyrGlnGlySerValArgGlnAsnSerArgProGlyLysProPhe 145
 DB 133 TTTCCTCACTCAAGATGCAAGAGCTCAGTAAGGAGAACTCTCGCCAGAAAGCCCTTC 74
 QY 146 LeuTyrValIleAsnAlaThrAspLeuAspProAlaThrProAsnGlyGlnLeuTyr 164
 DB 73 TTGTATGTCATGCGACAGACCTGATGATCCGCC-ACGCCCAATGCCAGCTTATAC 18
 RESULT 14
 BF334481 477 bp mRNA linear EST 22-NOV-2000
 LOCUS M01-CT0355-160200-002-g05 CT0355 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BF334481
 ACCESSION BF334481.1 GI:11305333
 VERSION EST.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 477)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR&ct=MR1-CT0355-
160200-002-905&ct3=2000-02-16&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 41.
Location/Qualifiers
1. 477
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0355"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT 160 a 104 c 95 g 118 t
ORIGIN

Alignment Scores:
Pred. No.: 8.09e-80 Length: 477
Score: 779.00 Matches: 151
Percent Similarity: 96.82% Conservative: 1
Best Local Similarity: 96.18% Mismatches: 5
Query Match: 18.03% Indels: 0
Gaps: 0

US-10-025-380-1081 (1-832) x BF334481 (1-477)

QY 312 AAlaAspGluTyrGlyProLeuSerTyrProLeuGluLeuHsValLysValLys 331
DB 6 GCATGAGATGAGACGAAACCACTTCAATCCCTGGAATACATGTAAGAGATAA 65
QY 332 AspIleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGlu 351
DB 66 AGATATTAATGATATCACTCACTCAATGTCCTCCACGATACCTATTGAGGTCCAGAG 125
QY 352 AsnGluArgLeuGluYanSerIleGlyThrLeuThrAlaHsAspArgAspGluGluAsn 371
DB 126 AATGAAAGCACTGGGTAAAGATATCGGACCTTACTGCATGACATGACGGATGAAGAAAT 185
QY 372 ThrAlaAsnSerPheLeuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAsp 391
DB 186 ACTGCCAAGATTTTAACTAAGAGATTTGAGACAACTCCAACTTCCCATGAT 245
QY 392 GlyLeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys 411
DB 246 GGAATCTTCTTAATCCAAACCTAAGCTGAGATGTACAGTAAACAGTCTTGAAG 305
QY 412 LysGlnAspThrProGlnTyrAsnLeuThrIleGluValSerAspLysAspPheLysThr 431
DB 306 AAGCAAGATATCTCTCACTAACAATGAGATGAGTGTCTGACAAAGATTCCAAGAC 365
QY 432 LeuGlyPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
DB 366 CTTGTTTGTGCAATCAACGTTATGATATCAATGATCAATGATCCCATCTTTGAAAAA 425
QY 452 SerAspTyrGlyAsnLeuThrLeuAlaGlyAspThrAsnIleGlySerThr 468
DB 426 TCGATTATGAAACCTGACTCTTGTCTGAAGACACACATGAGGTCCACC 476

RESULT 15

AM862312/c
LOCUS AM862312 563 bp mRNA linear EST 19-MAY-2000
DEFINITION RC4-CT0367-130200-011-b08 CT0367 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM862312
VERSION AM862312.1 GI:7958007
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1. (bases 1 to 563)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=RC4-CT0367-130
200-011-b08&ct3=2000-02-13&ct4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 499.
Location/Qualifiers
1. 563
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="CT0367"
/dev_stage="Adult"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
, 716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

FEATURES
Source

BASE COUNT 155 a 113 c 121 g 174 t
ORIGIN
Alignment Scores:
Pred. No.: 1.83e-79 Length: 563
Score: 777.00 Matches: 163
Percent Similarity: 92.97% Conservative: 9
Best Local Similarity: 88.11% Mismatches: 11
Query Match: 17.98% Indels: 4
Gaps: 0

US-10-025-380-1081 (1-832) x AM862312 (1-563)

QY 232 IleTyrValThrGlnProLeuAspArgIleGluLysAspAlaTyrValPheTyrAlaVal 311
DB 562 ATTATCGGACCTCACCCCTTGGACGAGAGATGATGATATATTTTATCAAT 503
QY 312 AAlaAspGluTyrGlyProLeuSerTyrProLeuGluIleHsValLysValLys 331
DB 502 CCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 443
QY 332 AspIleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGlu 351
DB 442 GAGATTATGATTAATCACTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 383

```
QY 352 AaenGIuArgIeunGIyaSenSerIleGIYThreuThraIaHIsaPaArgaSpGIu-As 371
Db 382 AATGTACGACTGGGTACAGTATCGGAGCCCTTACTGACATGACAGGATGAAGAAAG 323
QY 371 nTrAlaAsnSerPheLeuAnTYrArgIleValGIuGInThrProLYsLeuPromeIAs 391
Db 322 TACTGCCAACAGTTTCTATACAGAGATTGTGAGCAAACTCCAACTTCCATGGA 263
QY 391 pGIyleuPheLeuIleGInThrYrAlaGIyMetLeuGInLeuAlaLYsGInSerLeuY 411
Db 262 TGGACTCTTCTATCCGAACCTATGCTGGAATGTACAGTATACAGTCTTGAA 203
QY 411 sLYsGInAspThrProGInTYrAsnLeuThrIleGIuValSerAspLYsAspPheLYsTh 431
Db 202 GAAGCAAGATACCTCTCAGTACAACTTAACGATAGAGGTGTGACAAAGATTCAAGAC 143
QY 431 rLeuCySPheValGIuIleAsnVal-IleAspIleAsnAspGIuIleProIlePheGIuL 451
Db 142 CCTTTGTTTGTGCAATCAACGTTAGTATATCAATGATCAATCCCATCTTGAT 83
QY 451 ySserAspTYrGIyAsnLeuThreuAlaGIuAspThraSnIleGIySerThrIleLeuT 471
Db 82 A-TCAGATATAGTAACCTGACTCTGCTGAAGACACAAACATTGG--TCCACCATCGTAA 26
QY 471 hTrIleGInAla 474
Db 25 CCATCTTGCA 15
```

Search completed: June 22, 2003, 03:20:27
Job time : 2447 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 00:52:29 ; Search time 63 seconds
(without alignments)
2721.130 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321

Sequence: 1 MIQAHLSLCILMYLATG.....DKGKNVESQASVPLNS 832

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2332	54.0	578	11	Q63423
2	1427.5	33.0	868	13	Q90X63
3	709.5	16.4	893	13	Q80VQ7
4	699	16.2	922	13	P79883
5	677	15.7	783	13	Q90275
6	671.5	15.5	839	4	Q9HAZ9
7	652	15.1	713	4	Q8TXX3
8	652	15.1	714	11	Q8VDK4
9	651.5	15.1	814	6	Q77704
10	650	15.0	714	11	Q8R490
11	628	14.5	4589	11	Q9WU10
12	624	14.4	4587	11	Q90XA3
13	611	14.1	4555	11	Q8R508
14	601.5	13.9	3579	5	Q9VSN8
15	591.5	13.7	4643	5	Q9VW71
16	576	13.3	3014	4	Q9NY06

17	572.5	13.2	901	4	Q9UII7	Q9UII7 homo sapien
18	571	13.2	3034	11	Q35161	Q35161 mus musculus
19	570.5	13.2	840	4	Q9H800	Q9H800 homo sapien
20	570.5	13.2	894	4	Q9HB01	Q9HB01 homo sapien
21	564	13.1	3301	11	Q9IZ10	Q9IZ10 mus musculus
22	561.5	13.0	796	4	Q96CZ9	Q96CZ9 homo sapien
23	558	12.9	3313	11	Q98278	Q98278 rattus norv
24	557.5	12.9	3312	4	Q9NYQ7	Q9NYQ7 homo sapien
25	556.5	12.9	788	11	Q8V168	Q8V168 mus musculus
26	552	12.8	3503	5	Q24292	Q24292 drosophila
27	545.5	12.6	801	11	Q9Z0M3	Q9Z0M3 mus musculus
28	543.5	12.6	693	4	Q15066	Q15066 homo sapien
29	543.5	12.6	792	13	Q9DF80	Q9DF80 xenopus lae
30	543.5	12.6	798	13	Q8GCH3	Q8GCH3 gallus gall
31	541.5	12.5	792	13	Q9DF81	Q9DF81 xenopus lae
32	538	12.5	864	13	Q90Z37	Q90Z37 brachydanto
33	530	12.3	2923	4	Q9HC04	Q9HC04 homo sapien
34	528	12.3	2809	5	Q61230	Q61230 lytechinus
35	521	12.1	4349	4	Q9NYQ8	Q9NYQ8 homo sapien
36	518.5	12.0	4351	11	Q88277	Q88277 rattus norv
37	518	12.0	782	6	Q8NNW5	Q8NNW5 sus scrofa
38	515.5	11.9	910	5	Q95YK9	Q95YK9 citra savig
39	514.5	11.9	503	4	Q9NTY1	Q9NTY1 homo sapien
40	508	11.8	2920	11	Q9ROM0	Q9ROM0 mus musculus
41	496.5	11.5	790	13	Q91838	Q91838 xenopus lae
42	494	11.4	794	13	Q93264	Q93264 xenopus lae
43	489	11.3	828	4	Q9U199	Q9U199 homo sapien
44	489	11.3	1783	5	Q9VJB6	Q9VJB6 drosophila
45	488.5	11.3	821	4	Q9U118	Q9U118 homo sapien

ALIGNMENTS

RESULT 1

Q63423 PRELIMINARY; PRT; 578 AA.

AC Q63423;
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1996 (TREMBLrel. 21, Last annotation update)
DE Proton-driven peptide transporter (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NISTAR; TISSUE=SMALL INTESTINE;
RX MEDLINE=96067558; PubMed=7488096;
RA Erickson R.H., Gunn J.R. Jr., Lindstrom M.M., McKean D., Kim Y.S.;
RT "Regional expression and dietary regulation of rat small intestinal
peptide and amino acid transporter mRNAs.";
RT Biochem. Biophys. Res. Commun. 216:249-257(1995).
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL: I46874; AAC2077.1;
DR HSSP: P5116; INCI.
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin/6.
DR SMART: SM00112; CA; 6.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS50268; CADHERIN_2; 6.
KW Calcium-binding; Cell adhesion; Glycoprotein.
FT NON_TER
FT NON_TER
SQ SEQUENCE 578 AA; 64517 MW; 5644574BC6D14F80 CRC64;

Query Match 54.0%; Score 2332; DB 11; Length 578;
Best Local Similarity 77.3%; Pred No. 4.3e-145;

Matches 447; Conservative 55; Mismatches 76; Indels 0; Gaps 0;

QY 62 LTGETDNFVEREGILYNNRDLRETRSTNNLQVAALDANGIIVEGPVITREVDIND 121
||||| |::|||:| |||||:|::|||:| |||||

```

Db      1  LTGETDGLIKIEKQGLIYHARVLDRETRAVHHLRLAALDSQALVDPVPIIEVDIND 60
Qy      122  NRPTFLOSKTEGVRONSRPGKPLVYNATDLDPPATNGOLYYQIYIOLPMMNNWYFQ 181
Db      61  NRPTFLOSKTEGVRONSRPGKPLVYNATDLDPPATNGOLYYQIYIOLPMMNNWYFQ 120
Qy      182  INKKTGALSTRGSGOELNPAKNSYVLVSVKMGQSGNSFSDTTSVLIITENIMKA 241
Db      121  IDNKGTALSTRGSGOELNPAKNSYVLVSVKMGQSGNSFSDTTSVLIITENIMKA 180
Qy      242  PREVENENSTDPHPKITQVRWMDPGAOYSLVDEKELPRPPPSIDEGDIYVTOPLDRE 301
Db      181  PEVEIRENLTDPHPKITQVRWMDPGAOYSLVDEKELPRPPPSIDEGDIYVTOPLDRE 240
Qy      302  EKALYFVAVAKDEYKPLSYPLFIHYKVDINDNPTCSPTVFEVQENRGLNSIGT 361
Db      241  EKSHVFAFACQKENGKPLAYPLEIRVKVIDINDNPTCSPTVFEVQENRGLNSIGT 300
Qy      362  LTAHDEENSTANSFLAYRIVEOTPKLPMGDLFIQYAGMLQAKSLKKOPTPOYNLT 421
Db      301  PAHMDDEANNINSLMAYRLVDOTPKVPSDELFLIDYGGKVLGKSLKKODSPQYNLT 360
Qy      422  IEVSDKDFKLCFVQINVIDINDQIPIFEKSDYGNLTAEADTNIGSTILTIOATDADPE 481
Db      361  VEVSIDDFKLCGQVAVINDINDQIPIFEKSDYGNLTAEADTNIGSTILTIOATDADPE 420
Qy      482  TGSSTLIYHIIKQDSBGLGVDTDPHTNTGYIILKKLPDETAANSIVKAEENPELVF 541
Db      421  TGSSTLIYHIIKQDSBGLGVDTDPHTNTGYIILKKLPDETAANSIVKAEENPELVF 480
Qy      542  GVKYVNSPFAKFTLIYVDVNEAPQSOHVFOAKYSEDAVIGTKVANTADPEGLDYS 601
Db      481  GIEVNSPFAKFTLIYVDVNEAPQSOHVFOAKYSEDAVIGTKVANTADPEGLDYS 540
Qy      602  LRQDTRGMLKIDHYTGEIFSVAPLDRAGSEYRQVVA 639
Db      541  LKONKRGMLKIDHYTGEIFSVAPLDRAGSEYRQVVA 578

RESULT 2
ID 090X63  PRELIMINARY;  PRT;  868 AA.
AC 090X63;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Caderin-17.
GN CDH17.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;
RN [1]
SEQUENCE FROM N.A.
RA Hofffield J., Ramachandran A., Reuter K., Lavallie E.,
RA Collins-Racie L., Crostier K., Crostier P. St.;
RT "Zebrafish caderin-17 is required for proneuritic duct development.";
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 6 CADHERIN DOMAINS.
DR EMBL; AF428098; AAL29444.1;...
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 6.
DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_2.
DR PROSITE; PS50268; CADHERIN_2; 6.
KM Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 868 AA; 96677 MW; 40C30844C836E990 CRC64;

Query Match 33.0%; Score 1427.5; DB 13; Length 868;
Best Local Similarity 36.5%; Pred. No. 2.5e-85;
Matches 320; Conservative 150; Mismatches 330; Indels 77; Gaps 16;
13 LMLYIATGYG--QGKSGPLKMTFTSYEGQSFQIIFQF--KANPRAVTFELNGETDNI 69

```

```

Db      9  LTLVLSIGHIDLEBK--KGPLDITVDVEATVPVPAFAFKFTSAVEDVSSYVSGETEDK 67
Qy      70  FYIERGGLIYNRALDRETRSTHNLQVALDANGIIVVEPVTITTEVCHINDNRPFLDS 129
Db      68  IRISDGMVLYLOPLEWSEBPKHLLIEALSBDGKLDGPAAVYVQVINDVNNPVSFS 127
Qy      130  KYEGSVRONSRPGKPLVYNATDLDPPATNGOLYYQIYIOLPMMNNWYFQIINKTGAI 189
Db      128  QISGSIRESIPAGVFPVQVPSADDDPNTENTOLRSIYNQIPIVQGTFFGGINPNNGI 187
Qy      190  SLTRGSGOEL----- 199
Db      188  FTTEBGAFLKARPSVTYSRGEVRGSDVLKKKFEDECIPKNNIALENNPYKCYERAR 247
Qy      200  ---NPAKNSYVLVSVKMGQSGNSFSDTTSVLIITENIMKAAPVEMTENSNDPDP 256
Db      248  RTVNVLYQDDVDVALIYVAEDLGNAVNSLSTTRVNAIILQNLVWSFGPITIENTDEEP 307
Qy      257  IKITQVRWMDPGAOYSLVDEKELPRPPPSIDEGDIYVTOPLDREKDAYVFAVAKDEY 316
Db      308  MYLATFRANNPTALYRLQEKEL--SPTTINDGDIIYVTPDREKEMAYILVIAEDQG 366
Qy      317  GKPLSYPLFIHYKVDINDNPTCSPTVFEVQENRGLNSIGTITAHDRDEENTANSF 376
Db      367  GVELEKMEIPLYQDENNDPPMCD--ALFEVQEKEPVNGSIGHLPAHNDKEGTLSSA 424
Qy      377  LNRIVYEQPKLPMGDLFIQYAGMLQAKSLKKODTPQYNLTIEVSDKDFKLCFQ 436
Db      425  LVTYLSQTPPKSDGMFSDIDPNTGEIKVANOFPQVQYELFEVVDQVFFTKCAI 484
Qy      437  INVINDINDQIPIFEKSDYGNLTAEADTNIGSTILTIOATDADPEFGSSKILYHIKGS 496
Db      485  IKVIDINDQIPIFEKSDYGNLTAEADTNIGSTILTIOATDADPEFGSSKILYHIKGS 544
Qy      497  EGRGVDTDPHTNTGYIILKKLPDETAANSIVKAEENPELVFVKYVNSPFAKFTLI 556
Db      545  QNLFALVEDEBERGEGVYIAQPLDYELQSVNLIKIDARNEPLIAGVEYVNDSTSVIE 604
Qy      557  YVDVNEAPQSOHVFOAKYSEDAVIGTKVANTADPEGLDYSYLRQDTRGMLKIDHYT 616
Db      605  LVVDDEPPEVEEGLNVNVPENITVGTLLMKAEAKDPEKTIKFKNEGDEHKLELVDT 664
Qy      617  GEIFSVAPLDRAGSEYRQVVAATEVYGSSLSVSEFHLIMDVNDNPRILAKDYGLF 676
Db      665  GELKTAALDRETVDFHTLTITRYETSGKMEAMKVDIHLQVDVNNYFKLQK--TQGR 722
Qy      677  CHPLSAPSLIFEATDDOHLFRGPHFTFSLSGSLQNDMEVSKINGTHARLSTRHTDE 736
Db      723  CLQDMP--LTLAMDKADPY--GEPTFAISRKS--QNFRIKPVDTGSAKILKKKPS 777
Qy      737  ERAYVVLIRINDGRRPLBGIV--SLVYTCSCYE--GSCF--RPRGHGTGPTVMANGILL 793
Db      778  EQNVTVPIVNLDAAG--LGIQKFPVRICTKTKGVCYIIPASHMKWL--SMGSTIGILA 833
Qy      794  TLLVYGIILAVFIRIKKDKQDNVESQAQSEVKPL 830
Db      834  GVFGVIGLFGICLVQIKK--KOKORATAREGTTKAM 867

RESULT 3
ID 08UV07  PRELIMINARY;  PRT;  893 AA.
AC 08UV07;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE N-cadherin.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_Taxid=7955;

```

```

RN [1]
SEQUENCE FROM N.A.
RA Lele Z., Folchert A., Concha M., Rauch G.-J., Geisler R., Rosa F.,
RT Wilson S.W., Bailly-Maître L., Hammerhead M.;
RT "zebrafish paracrine/n-cadherin is required for neural tube closure
RT and correct positioning of neurons within the CNS.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF418565; AAL59160.1; -
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR InterPro; IPR000215; Serpin.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_3.
DR PROSITE; PS50268; CADHERIN_2; 3.
DR PROSITE; PS00284; SERPIN; UNKNOWN_1.
SQ SEQUENCE 893 AA; 98990 MW; 5E3FF0CDD92B3A6F CRC64;

Query Match 16.4%; Score 709.5; DB 13; Length 893;
Best Local Similarity 27.8%; Pred. No. 4,5e-38;
Matches 213; Conservative 115; Mismatches 335; Indels 103; Gaps 21;

QY 123 RPTFLQSKTEGSRKNSRPGKPLVYNADL-----DDPATNGQLYQIVQLP 172
DB 28 QPGFTNEVYVMTADVITTEGQVLKVDVDCGRSGSLRFESGDPAD----- 73
QY 173 MINWVYFOINNTGALISLRSGSELNPAKNSYLVISVKGMSQSE-----NSFSDT 228
DB 74 -----FRI-DADGVTMAAR--TLQTLDRKGS--LEIKAKENSQEQMWHINFTQPK 121
QY 229 SVDIIV-----TENIWKAPKPEVENSSTDPH---IKITQVRMNDPGA 269
DB 122 QVPVILFPRHSVLVKGDSDVNRKRDMVLP--PNNVLENSRKOPPEELVKIQSXKSNLT 180
QY 270 QYSLV--DKELRPPPSIDQ--EGDIYVTPQLDREKDAVYFAVAKDEYKPLSYPLR 326
DB 181 RYSVTGPAGADQNTGFIIDPIISGLSVTKPLDREH1PWFHLAAHVDINGNQMEMPIDI 240
QY 327 HYKVVQDINNPRTCPSPVTVFEQENBRKNSIGTLTADRDENRANFLNRYIEQRP 386
DB 241 IINVDIMNNRPFETHQINNGVTDSEAKPFTVITSDCKDPNTANGELRKILISQTP 300
QY 387 KLPMDLFLIQTAYAGMLQAKOSLKKQDTPQVNLTEVSDK-----FKTLCFVQINYI 440
DB 301 EGSSSSMFTINNKTKGILITVAAGLDREKYPQYTLIIQADMEENPFGYGSNATAVIRLI 360
QY 441 DINDQIPTEKSDYGNLTLAEDTNIGSTLITQATDADEPFGSSKILYHIKDSGRL 500
DB 361 DVNDNAPEFTRETFHGEVEENRNVNIVTLTV--TDKDEPFPANMAVYRIISGDPYGRF 418
QY 501 GVDTPHTNTGVYIIKKPLDFTFAVNSNIVFKANEREPVFGYKYNASSPAKFTLIVTV 560
DB 419 STPLDVTNIEGLVTVKVPDFENNSFMLTVADNEVPLASGHRTRQSTAVVSIRVIV 478
QY 561 NEAPQSQHVPOAKYSEDAVIGTKVGNATKADPEGL---DISYSLGDRGMKIDHYTG 617
DB 479 NSPNNDPRPKQIKLEEGLPQWSMLTTFNHPDRMOOTISYSKLYDPAAMLEIDPNNG 538
QY 618 ELPFSAVPLDREA---GSPYVQVAVTEVGSSLSVSEFHLIMDVNDNPRLLADYTG 673
DB 539 RISTIAVLDRESPYKGNLYNATFMASDNGVPRASTGTGLQIYLLDINDNARV----- 592
QY 674 LFFCHPLSA-----FGLSLEFATDDQHLFRGHFTSLGS--GSLQNDMEVSKINT 724
DB 593 -----FQSEVCEEREPNAINITAVDGINPNAGP--YAFELPNRPSDRIIRMTLIRISD 647
QY 725 HARLSTRHDFEERAYVVLIRINDGRPLLEGIVSLPTVFCSC--VEGSCFRPAGHOTGP 783
DB 648 HAQIGIKISYLSGIEYELFISITDSGNLPMSTYTIKIRIVCQCDHNGDCV---DMERIM 703
QY 784 TVGMAVGILLTLLVIGIILAVVIRIKKDKNDVESQAASEVPR 829

```

```

DB 704 AAGTGAITAIILICIIILVLMFVMMKKRDKERQAKQLIDP 749
RESULT 4
ID P79883 PRELIMINARY; PRT; 922 AA.
AC P79883;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Maternally expressed neural cadherin Xmn-cadherin.
GN Xmn-CADHERIN.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodidae.
OX NCBI_TaxID=8353;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=TAIL BUD;
RC MEDLINE=96257964; PubMed=8652409;
RA Tashiro K., Tooi O., Nakamura H., Koga C., Ito Y., Hiyasa H.,
RA Shiohara K.;
RT "Cloning and expression studies of cDNA for a novel Xenopus cadherin
RT (Xmn-cadherin), expressed maternally and later neural-specifically in
RT embryogenesis.";
RL Mech. Dev. 54:161-171 (1996).
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
CC SORTING OF HETEROGENEOUS CELL TYPES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE 1 MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 3 CADHERIN DOMAINS.
DR EMBL; S62457; AAB37685.2; -.
DR HSSP; P15116; INCI.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin_5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS50268; CADHERIN_2; 3.
KW Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 922 AA; 101173 MW; FEA7A95CBAF1B640 CRC64;

Query Match 16.2%; Score 699; DB 13; Length 922;
Best Local Similarity 26.4%; Pred. No. 2,3e-37;
Matches 224; Conservative 140; Mismatches 334; Indels 152; Gaps 31;

QY 5 AHHSLCLIMLYLANTGYQDEKFGSLKPMFTSIEGQESQIIFQ--FRANPAVTEFLT 63
DB 34 ASVSSAC-----TPGFGADG--YTLVSP--NIEGQKLVKVKNGSSGAGIWE--T 82
QY 64 GETDNIPIYIEREGILLYRALDRETRSTHNLQVALDNGIIV---EGPPIITEVNDI 119
DB 83 NNPD--FVYAGDGAIV---TAREVQIPKQKFTVAANDHTP----- 120
QY 120 NDNRPFLQSKYEGSRKNSRPGKPLVYNATDLDPAVPGQLYYQIVQLPMINVMY 179
DB 121 -----EKWEALAIQ-----LFVE-----ETPLNQSQYQ----- 142
QY 180 FOINNTGALISLRSGSE-----LNPAPKPSNLYVSVKMGQSNPSDTSVDIT 233
DB 143 -----TG--SQSQSEQSQSQSGTLFWRHNGKLRKQKD----- 176
QY 224 VTENIWKAPKPEVENSSTDPHPIKITQVRMN---DPAQVSL--YDKEKLPRPESTDQ 288
DB 177 -----WVLP--PNNVLENSRGPPOQVLIRSDKRDITRISITGVGADQPMIAFNIDP 230
QY 289 -EGDIYVTPQLDREKDAVYFAVAKDEYKPLSYPLRHYKVVQDINNPRTCPSPVTVF 347

```

Db 231 IEGRAWVTRPLDEBESSYHLRAHVD INGNKVENPIDLSIYVDNDNRPFFSPFNG 290
 Qy 348 EVOGENRLGNSIGTLTAHDEBENTANSFLNRIVEQTPRLPMDGLFLITGAGMLQAK 407
 Db 291 SYDEAKSPGYVMTVAHADADDINTSGIWMYRIMOQSPSPSHDFVHSKTVGINTVA 350
 Qy 408 QSLAKDTPQYNLTIVSDKD-----FKTLCFVQVINVINDINOQIFPEKSDGNLTAE 461
 Db 351 AGIDREKVOQYVIVTQATDMEGNLNGLSNTAATVAVDNDNRPFFTKMFE--IGEPV 408
 Qy 462 DTNIGSTIITLQATDDEPFGSSKILYHIIKDSBGLVDPDPTHTNGYVILIKKPE 521
 Db 409 ENHVDVAVNLTVYVDDQPTSMNNAVKIISDPGHFRTKDPVLTNEGIVVSKPVYD 468
 Qy 522 ETAAVSNIYFKAENPEPELVFGVKNASFAKFTLITVDNEAPQFQCHVQAKVSEDAV 561
 Db 469 EMKSVFPLIWMVWNOAPLASGIOMSLQSTAAVTVSVNDVAVDFPRNPEIRKLEESA 528
 Qy 582 GTKVGNTAKDPE---GLDISLARGDTRGMLKIDHTGELFSAVAPLDEAGSP----- 632
 Db 529 GRLLITPSAVDPHSMQOVLARKISIDPANMLAINTNGVSTTAVLDRE--SPFYKDL 586
 Qy 633 YRVQVATEVGGSSLSVSEFHLILMDVNDNPPR-TAKDYTGLFCHPLSAPSLFEAT 691
 Db 587 YQAKFLATDNGNPPASGTGLLIQIDINDNAPELLPKDAQ---IGERPNGG-INITAL 642
 Qy 692 DDDQHLFRGPHFTPLSGS--GSLQNDWEVSKINGTARLSRTHTDPEERAYVYLIRNDG 749
 Db 643 DVDRKPSADP-FVFEPLPSVPTTIRRMWTIRHINSIYARLSIQGYLESQWYDVAVITDS 701
 Qy 750 GRPELBEIVSLPTFGSCVSGSCFRPAGHOTGIPVGMVAGILLTLLVIGITLAV--P 807
 Db 702 GNPPLNTSITIKKVCPCDNNGDCTTG---AVALAGLGAIIISILCIITLISVLLP 758
 Qy 808 IRIKDKGKD 817
 Db 759 VMMKREKE 768
 RESULT 5
 090275 PRELIMINARY; PRT; 763 AA.
 AC 090275;
 Dt 01-NOV-1996 (Tremblrel. 01, Created)
 Dt 01-NOV-1996 (Tremblrel. 01, last sequence update)
 Dt 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Neural-cadherin precursor (N-cadherin).
 OS CDH2 OR CDH2 OR ZNCAD.
 OS Brachydanio rerio (Zebrafish) (Zebra danio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_Taxid=7955;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=EMBRYO;
 RX MEDLINE=95178741; PubMed=7873785;
 RA Bitzur S., Kam Z., Geiger B.;
 RT "Structure and distribution of N-cadherin in developing zebrafish
 RT embryos: morphogenetic effects of ectopic over-expression";
 RT Dev. Dyn. 201;121-136(1994).
 RL Dev.
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PRESENTATIONALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN
 CC NEURONAL RECOGNITION MECHANISM.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY)..
 CC -1- TISSUE SPECIFICITY: EXPRESSION IN THE EMBRYO IS INITIALLY FOUND IN
 CC ALL DEEP CELLS BUT LATER BECOMES RESTRICTED TO VARIOUS EPITHELIAL
 CC AND NEURONAL TISSUES. FOUND IN DISCRETE AREAS OF CELL-CELL
 CC ADHESION WHEN EMBRYO IS NEARLY COMPLETED. EXPRESSED IN ADULT
 CC BRAIN, EYES AND TRUNK.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS FIRST DETECTED IN THE EMBRYO

CC DURING GASTRULATION.
 CC -1- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
 DR EMBL; X67648; CAA47890.1; -.
 DR HSSP; P15116; INCI.
 DR ZFIN; ZDB-GENE-990415-171; cdh2.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C_term.
 DR Pfam; PF00028; cadherin; 5.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 5.
 DR PROSITE; PS00232; CADHERIN 1; 2.
 DR PROSITE; PS0268; CADHERIN 2; 5.
 KW Cell adhesion; Glycoprotein; Phosphorylation; Transmembrane;
 KW Calcium-binding; Repeat; Signal.
 FT SIGNAL 1
 FT PROPEP 35
 FT CHAIN 36
 FT DOMAIN 36
 FT TRANSMEM 602
 FT DOMAIN 623
 FT REPEAT 36
 FT REPEAT 144
 FT REPEAT 259
 FT REPEAT 374
 FT REPEAT 481
 FT DOMAIN 740
 FT CARBOHYD 149
 FT CARBOHYD 201
 FT CARBOHYD 278
 FT CARBOHYD 331
 FT CARBOHYD 449
 FT CARBOHYD 499
 FT CARBOHYD 528
 FT CARBOHYD 569
 SQ SEQUENCE 783 AA; 87120 NM; B57ECBDP41CCIC5 CRC64;
 Query Match 15.7%; Score 677; DB 13; Length 783;
 Best Local Similarity 30.1%; Pred. No. 56-36;
 Matches 187; Conservative 98; Mismatches 288; Indels 48; Gaps 15;
 Qy 239 WKAPKEVENENSTDPHP---IKITQVRWMDGQOYSLV-DKEKLPRFPESIDQ-EGDI 292
 Db 37 WVLP-PVNVLENSRKQPEELVKIQSDKXKNTLRKSVTGPADQNPGLFIIDPISGL 95
 Qy 293 YTTQPLDREKDAYVYVAKBYGKPLSYPLEIHKKYKDINDNPTGCSPTVFEVQGN 352
 Db 96 SVTKPLDRHIFPHLRAHVDINGNQMENPIDIINVIDMNDNRPFFTHQIWNQTVDS 155
 Qy 353 ERLGNSIGTLTAHDEBENTANSFLNRIVEQTPRLPMDGLFLITGAGMLQAKSLK 412
 Db 156 AKFGTVMVTISQDKDPPITANGMLRYKLSQTPESPSSMFTINNKTKIITVAAGLR 215
 Qy 413 QDTPQYNLTIVSDKD-----FKTLCFVQVINVINDINOQIFPEKSDGNLTAEPTNG 466
 Db 216 EKVPOYTLITQATDMEGNPTVGLSNTATATVILVDNDNAPERTRETFGEVBNENVA 275
 Qy 467 STLTITQATDDEPFGSSKILYHIIKDSBGLVDPDPTHTNGYVILIKKPELDETA 526
 Db 276 VTMLTV--TDKDEPGPANNAVRIISGPTGFSIPTPVNTEGLVTVKVPDFEMNS 333
 Qy 527 SNIVFAENPEPELVFGVKNASFAKFTLITVDNEAPQFQCHVQAKVSEDAVISTKY 586
 Db 334 EMLTVVADNEVPLASGIHRTOSTAVSRVIDNESPNDPRPKIKLEBGLPQMSMLT 393
 Qy 587 NVTA-KDEBGL---DISYSLRGDTRGMLKIDHTGELFSAVAPLDEA---GSPYVQV 638
 Db 394 TFRHDDPDKYMQOTISYSLVDPAVLEIDPNNGRISTAVDRSPYKNNLYAVATEM 453
 Qy 639 ATVGSSSLSSVSEFHLILMDVNDNPPRLAKDYTGLFCHPLSA-----PGLSLEAT 691
 Db 454 ASDNGVPASGCTGLIYLDINDNAPRV-----FPQBAVCERPEPNAINITAV 503

```

QY 692 DDDQHLFRGPHFTFSGS--GSLQNDMEVSKINGTHARLSRTHDPEERAYVVLIRINDG 749
DB 504 DGLININAGP-YAFELPNRPSDIRRWLTLRISGDHQLSLKISYLSGIELPISITDS 562
QY 750 GRPLLEGIYSLPTVFGSC-VEGSCFRPAGHOTGIPYGVMAVGLLTLVLVIGIILAVFR 808
DB 563 GNLPENNTYLRKIKVQCQDHGDCV---DMERIMAGLGTGAILIICITIIILLVLVLM 618
QY 809 RIKKDKKDNVESAQASEVXP 829
DB 619 FVMMMRBRDKEKRAKQLIDP 639

RESULT 6
O9HAZ9 PRELIMINARY; PRT; 839 AA.
ID O9HAZ9;
AC O9HAZ9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Desmocollin 3b.
GN DSC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20484176; PubMed=11027496;
RA Whitlock N.V., Hunt D.M., Rickman U., Malhi S., Vogazianou A.P.,
RA Dawson L.F., Eady R.A., Buxton R.S., McGrath J.A.;
RT "Genomic organization and amplification of the human desmosomal
RT cadherin genes DSC1 and DSC3, encoding desmocollin types 1 and 3.";
RT Biochem. Biophys. Res. Commun. 276:454-460 (2000).
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
DR EMBL; AF293359; AAG23427.1; -.
DR HSSP; P15116; INCU.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 5.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 5.
DR Calcium-binding; Cell adhesion; Glycoprotein.
SQ SEQUENCE 839 AA; 93497 MW; F21BE2B865A6855 CRC64;

Query Match 15.5%; Score 671.5; DB 4; Length 839;
Best Local Similarity 30.8%; Pred. No. 1.3e-35;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

QY 239 WKAPKVEVENSTDPPIKITQVRWNNDGQAYSL-----YDEKELPRFPESIDQSG 290
DB 137 W-APICSGQENSLGFPFLQGV-SDAQNTVTVYSISRGVDKPEPLNLFIERD-TG 193
QY 291 DIIVTQPLDREEDKAY--VFYAVAKDEYKGPLSYPLEIHVKVDINDNPTCPSPYTV 348
DB 194 NIECTRPVREEDVDLAAYASTADYSADL--PLPLRVEDENDNHPVEFTEALYNE 251
QY 349 VQENELGNSIGTLTHADREENTANSFLNRYIVEOTPKLPMDGLFIQTAYMLQLANQ 408
DB 252 VESSRPGTGVVCAITDDEPDMTRKYSILQOTPRSP--GLPSVSPSTGVTITVSH 309
QY 409 SLKQOTPOYNLTIEVSDQFK-----TLGVOQNVINDINDQIIFEKSDQGNLTAE 461
DB 310 YLDREVVDKSLMKQDMQGFGLIGTSTC--ITVDSNNATPFQONAE--APVE 365
QY 462 DTNIGSTIIITQATDADEPFTGSKLTHIINKDSEGRGLVDPDPTNTGVYIIKKPLDF 521
DB 366 EAAFNVEILRIPIEDDLINTAMWRVNFLLKGENGHFKISTDKKTNGSVLSVXPLMY 425
QY 522 ETRAASNIYFKAENPEPLVFGV-KYNAASFATLTVTVTNEAPQPSQHVPAQKSEVYA 580
DB 426 EENROYNLEIGVNNAPFARDIPRVATLNRALVTVAHRLDDEGECTPAQYVRIKENLA 485

```

```

QY 581 IGTKVGNTAKDPE---GLDISYLRGDTRGMLKIDHVTGEIFSVAPLDRAGSP-----Y 633
DB 486 VCSKINGYKAYDPENRNGNGGLRYKKLHDPKGNTTDEISGSIITSILDRVETPNEXY 545
QY 634 RQOVVATEVGSGLSSSEFFHLLMDVNDNPRLAADYGLFGRCHPLSAPGLIFPATD 693
DB 546 NITVLAIIDKDDSCCTGLAVN--IEDVNDNPEIIQEX--VVICRKMGYTDIL--AVDP 599
QY 694 DQHLFRGPHFTFSGS--LQNDMEVSKINGTHARLS--TRHDPERAYVVLIRINDG 750
DB 600 DEPVHGAQ-FYFSLPSTSPSIRLSLTKVNDPAALSTQKNAQGE--YTIPIYKD-- 654
QY 751 RPLLEGIYSLPTVFGSCVVEGSCFRPAGHOTGIPYGVMAV-GILLTLVLVIGIILAVFR 809
DB 655 RAGQATKLRNLNLCCHTPTQCRATSRSTGYILGKMAILALLIGALLFSVLLTVCGV 714
QY 810 IKKDKK 816
DB 715 FGATYK 721

RESULT 7
O8TBX3 PRELIMINARY; PRT; 713 AA.
ID O8TBX3;
AC O8TBX3;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Cadherin 13, H-cadherin (heart).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-TESTIS;
RA Strausberg R.;
RT Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC028624; AAH28624.1; -.
SQ SEQUENCE 713 AA; 78320 MW; E947F5FA773988FA CRC64;

Query Match 15.1%; Score 652; DB 4; Length 713;
Best Local Similarity 28.9%; Pred. No. 1.9e-34;
Matches 192; Conservative 100; Mismatches 274; Indels 98; Gaps 21;

QY 160 NQOLYQIVITQPMINNVYFQINKTGAISL---TRESGQLNPAKNPSYN-----LVI 211
DB 61 NDKLRYEV-----SSPYFVNSDGLVLRNITAVGKTLFVHARTPHADMAELVIV 112
QY 212 SYKDMGGQSEN--SFSDTTSV-----DIIVENIKAKRPVEMVENSNDPHIKITQVRM 264
DB 113 GSKDQGSLODIFKFAKTSVPRQKRSIVS-----PILIPENQRPFRDVGK-- 162
QY 265 NDPGAQYSLVDEKELPRFP-----SIDQ-----GDIYVQPLDREEDKAYVF 308
DB 163 -----VSDRREKSRFRLTGKGVDEPBGIFRINENTSVMTRTLDEAVAYQL 213
QY 309 YVAVDEYKPLSYPLEIHVKVDINDNPTCPSPVTVFEVQENRNGNSIGTLTHADRD 368
DB 214 FVEITDVNGKTLGEPVLEIVIIDQNDNRPIFREPGYIGHVWEGSPGTGVVNRMTAFDAD 273
QY 369 EENTANSPLNRYIVEOTPKLPMDGLFIQTAYG-MQLAKGSLKQDPT---POYNLTIV 424
DB 274 DATNALLRNIRIQTPDKPSPNMFYIDPEKGVITVAVSPALLDRELENKYEYELIEA 333
QY 425 SDK-----DFKTLCFVQNVINDINDQIIFEKSDQGNLTAEVNIITQATDADE 479
DB 334 QMAGDGLGCTATATNTIMDKNDHSKFTYKGE--QATVEGAVG--YIVVLVYEDKDD 390
QY 480 PFTGSSKLIYHIINKDSEGRGLVDPDPTNTGVYIIKKPLDEETAASNIYKCANPEPL 539
DB 391 PATGAMRAAYTIIINGNPGOSFEIHTNPQTNBGLSVVPLDYEISAFTLLIKVENEDPL 450

```

QY 540 VEGVKNASSFAKTLIVDVNEAPQESQHVFOAKVSEDAVIGTKVGNATKADPEGLD-- 597
 DB 451 VPDVSTGSPSTATVHTITVDVNGVFPVPPDMVTTQEDLSQESVLLTVNATPDSLQHO 510
 QY 598 -ISYSRGDTRGMKIDVHGEIFSVAPLDREA----GSPRYQVAVATEVGGSSLSVSE 652
 DB 511 TIRYSYKDPAGMLNINPINGVDTTAVLDRESTFVONSYYTALFLAIDSGNPPATGTGT 510
 QY 653 FHLIMDVNDNPPRLAKDVTGLFCHPLSAP-----GSLIFEATDDOHLFRGPHT 704
 DB 571 LITLEDVNDNAP-----FIYPTVAECDDAKNLSVILGASDKDLHPNTDP-FK 619
 QY 705 FSLGSSGLND-WEVSKINGTHARLSRHTDFEERAVVILRINDGRPPLEGIVSLPVT 763
 DB 620 FEIHKQAVDPKWKISKINTHALVSLQ--NLNKANTNPLIMVTDGSKPMTNITDLRVO 678
 QY 764 FCSC 767
 DB 679 VCSC 682

RESULT 8

OBVDK4 PRELIMINARY; PRT; 714 AA.
 AC 08VDK4;
 DT 01-MAR-2002 (TEMBLrel. 20, Created)
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Caderlin 13.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC021628; AAH21628.1; -
 DR InterPro; IPR002126; Caderlin.
 DR Pfam; PF00028; caderlin.5.
 DR PRINTS; PRO0205; CADHERIN.
 DR SMART; SM00112; CA: 5.
 DR PROSITE; PS00232; CADHERIN_1; UNKNOWN_3.
 DR PROSITE; PS50268; CADHERIN_2; 3.
 DR PROSITE; PS50268; CADHERIN_2; 3.
 SQ SEQUENCE 714 AA; 78116 MW; CASCA791C46A2FFB CRC64;

Query Match 15.1%; Score 652; DB 11; Length 714;
 Best Local Similarity 29.3%; Pred. No. 1.9e-34;
 Matches 189; Conservative 109; Mismatches 287; Indels 60; Gaps 20;

QY 160 NGOLYQVIVQLPMINNVMTFOJNNKTGAISL--TEGSOELNPAKPNSTN----LVI 211
 DB 61 NEKLMHYEV-----SSPHFKVNSDGLVALRNITAVGRTLFVHARTPHADMALVIV 112
 QY 212 SVYDMGQSEN--SFSPTTSV-----DIYVENWKAAPKPEWENSTDDHPKINOVW 264
 DB 113 GGDIDIGSLQDIFKFAFRTSPVPROKRSIVS-----PLIENOROPFRDVGAVD 164
 QY 265 ND--PGAQVSLVND--EKLPRFPFSIDOE--GDIYVTOPLDREKDAVYFAVADKEXKP 319
 DB 165 SDREGSKFRLTGQVQODPKGTFRINENGSVSTLTLDRETITATQLYVETTDASGT 224
 QY 320 LSYPLEIHVKVDINDNPPCTSPVTVFEVQENRNLGNSICTLTAHDEENTANSLY 379
 DB 225 LEGVPLEVIVYIDNDNRPIFRGPRYIGHVEGSPGTGTWRMTAFADDPATDNALIRY 284
 QY 380 RIVQOTKRLPMDGFLIQTIVG-MLQAKOSLKKQDT---PQNLITEVSGK-----DFK 430
 DB 285 NIRQOTDKPSPPNMFYIDPEKGDIVTVSSPLLDRETLNKKYLLIEAQMACLDVGLT 344
 QY 431 TLCEVQIVNDINDQIPFEKSDVGNLTADTNIGSTILTITQATDDEPTGSKILYH 490

DB 345 GTATATVIDDKNDSPKFTKEF--QATVEGAVG-VIVNLTVEDKODPTTGAMRAAYT 401
 QY 491 IIKGDSSEGSLGVDTPHTNTGYIILKKPLDFTAAVSNIVFAKENDPEPLVFGKYNASSF 550
 DB 402 IINGNPGOSFEIHNQVQNEGMLSVKPLDVIISAFTILLIKVENBDPLVPDVSTGSPST 461
 QY 551 AKFTLVTDVNEAPQESQHVFOAKVSEDAVIGTKVGNATKADPEGLD--ISYSRGDTR 607
 DB 462 ATVHTITVDVNGVFPVPPDMVTTQEDLSQESVLLTVNATPDSLQHOQTIRESYKDBA 521
 QY 608 GMLKIDVHGEIFSVAPLDREA----GSPRYQVAVATEVGGSSLSVSEFHLIMDVNDN 663
 DB 522 GMLSINPINGVDTTAVLDRESPFVNSYYTALFLAIDSGNPPATGTGTLITLEDINON 581
 QY 664 PRLAKDVTGLFCHPLSAPGSLIFEATDDOHLFRGPHTFSLGSSGLND-WEVSKIN 722
 DB 582 APVITP--TVAECDDARNLVILGASDKDLHPNTDP-FKEIHKQVTPDKWKISKIN 638
 QY 723 GTHARLSRHTDFEERAVVILRINDGRPPLEGIVSLPTECSC 767
 DB 639 NTHALVSLQ--NLNKANTNPLIMVTDGSKPMTNITDLRVOVCSC 682

RESULT 9

OBVDK4 PRELIMINARY; PRT; 814 AA.
 AC 077704;
 DT 01-NOV-1998 (TEMBLrel. 08, Created)
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Desmocollin type 2 (Fragment).
 OS DSC2.
 OC Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98380123; PubMed=9716266;
 RA Roberts G.A., Burdett I.D., Pidgeley S.C., King I.A., Magee A.I.,
 RA Buxton R.S.;
 RT "Antisense expression of a desmocollin gene in MDCK cells alters
 RT desmosome plaque assembly but does not affect desmoglein expression.";
 RL Eur. J. Cell Biol. 76:192-203(1998).
 CC -/- SIMILARITY: CONTAINS 4 CADHERIN DOMAINS.
 DR EMBL; A0002299; CAA03509.1; -
 DR HSSP; P09803; ISUH.
 DR InterPro; IPR002126; Caderlin.
 DR Pfam; PF00028; caderlin.5.
 DR SMART; SM00112; CA: 4.
 DR PROSITE; PS00232; CADHERIN_1; 2.
 DR PROSITE; PS50268; CADHERIN_2; 4.
 KW Calcium-binding; Cell adhesion; Glycoprotein.
 FT NON TER 1
 FT CHAIN 101 814
 SQ SEQUENCE 814 AA; 91081 MW; 59EF086C69BD3882 CRC64;

Query Match 15.1%; Score 651.5; DB 6; Length 814;
 Best Local Similarity 30.2%; Pred. No. 2.5e-34;
 Matches 181; Conservative 110; Mismatches 256; Indels 53; Gaps 20;

QY 239 WKAPKPEWENSTDDHPKIKITQVRWNDGCAQYSL-----VDREKLPFRFPFSIDOE- 289
 DB 102 W-APICSMQENSLGFPFLFOIQ--SDTANQNTIYISIRGPVDRE--PKNLFYVERDT 157
 QY 290 GDIYVTOPLDREKDAVYFAVADKEXKPLSYPLEIHVKVDINDNPPCTSPVTVFEV 349
 DB 158 GNLFCRPPVDREYSEFELIAFATPDGTYBELPLVLVIRIBEDNDNVPITFEKTYVFTI 217
 QY 350 QENRNLGNSICTLTAHDEENTANSLYRIVYEQPKLPMQDLFLIQTAYGMLQAKOS 409
 DB 218 SENCRGSTVGVQVCAVDKDEPDMTRLKXSIIEQPAVFT--LFSMHPATGVITTSQ 275


```

Oy 410 LKKOQTPQVNLITLVEBDXO-----FTKFLGVQVQNVINDINDQIPFESQSDGNLTLMEDTN 464
Db 276 LDRELDIDKQULKIKQMDMGQYFGLOTTSCITCINIDVDNDNLPTFRITST--VTSVEANT 333
Oy 465 IGSITLLITQATDADBEPTGSSKILYHIKNGSEBGLVDPTDPHTNGYVILIKKPLDPEYA 524
Db 334 VDVEILRTVTEVDKDLMTANMFRANYITLNGNENENFKIYTDPTNGSILLCVAFPLVNEER 393
Oy 525 AVSNIVFKAEENEPL-VFEGVKTNASSFAKFTLIVTDVNEAPQFSOHVFOAKSEDAVAGT 563
Db 394 KOVDLOIVVNNAPYSEKESLSTWSTAVTVVNNQODBEGPESCPVQIOTIOTENPVGT 453
Oy 584 KVGANTYADQPE---GLDISYSLRGDTRGMUKLIDHTYGEIJSVAPLDREA-----GSPRYQ 636
Db 454 EIQGYKAYDPERRRSSGIRYKKULSDPKDEKWRINENGTSTTFENLIDREAMIRGILYNIT 513
Oy 637 VVATEVGGSSLSVSSEFHLILMDVNNPNPRLADYTGTFCHPLSPASGLIFPATDDQH 696
Db 514 ILASDKDGRSCGT--LGITLQDVNDNGSIPKH--TVYICKTVLSPSADIV-AIDBEP 568
Oy 697 LFRGHPFTFSLGS---GSLQNDMEVSKINGTARLSRTHDPEER-----AYVVLIRIND 748
Db 569 I-NGBPPEPDRLENIIPDSDIERTRLTKINDTARL-----FYEKDLFPGTYSIRSVAD 621
Oy 749 GGRPFLLEGVLSVPYTFCSQV-EGSCRPAGHQIGIPIYGVMAVGILLTLVLIGILLAVF 807
Db 622 --RHORSILTPKVLVLCDCVTENDCTLRTARDARGGADVRLGKXVAILIALLIGILLFCLIF 679

```

RESULT	10		
QBR490			
ID	QBR490	PRELIMINARY;	FRT; 714 AA.
1C	QBR490:		
DT	01-JUN-2002	(TREMblrel. 21, Created)	
DT	01-JUN-2002	(TREMblrel. 21, Last sequence update)	
DT	01-JUN-2002	(TREMblrel. 21, Last annotation update)	
DE	T-cadherin.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=HEART;		
RA	Niermann T., Schmutz S.;		
RT	"Cloning of rat T-cadherin."		
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.		
QZ	EMBL; AF494095; AA014607.1; -		
QZ	SEQUENCE 714 AA; 78085 MW; 785614281625F864 CRC64;		

	Query Match	15.0%;	Score 650;	DB 11;	length 714;
	Best Local Similarity	29.1%;	Pred. No. 2,6e-34;		
	Matches	188;	Conservative 110;	Mismatches 287;	Indels 60; Gaps 20;
Qy	160	NGOLTYIVITVPLMNNWTFQINNKGTGAI---	SITREGSGELNPAKNPSTN----	LVI	211
Db	61	NKKAHYE-----SSPYFKNSDGLTAVNRNTAAGRILTFVARPRPHADEMAETIV			112
Qy	212	SVKMGGQSEN--SFSDTTSV---DIITENIKAKPEPVEMVENSIDPHIKITQVRW			264
Db	113	GKKIQGSLDQIFKARTSPYRQRSLTVS-----PIIIPNQGPFRPDVGKYVD			164
Qy	265	ND--PGAOISLVDR-EKLPPRPSSIDOE-GDIYTTPLDLNEKDAAVFVAVADGEYKP			319
Db	165	SDREPGSKFRLTGKGVDDPKGFTEFINENTGSVSATRLDRRTATAYDLFEITDASGR			224
Qy	320	LSYPLEIHAKVKDINDNPPTCSPTVVEVOENERLGISIGTLTAHRRDEENTANSFLNY			379
Db	225	LEGAPLEIVITVDQNDNRPRIREFGYIGHWEGSPGTGWRMKFAFDADDPATIDALLRY			284
Qy	380	RIVERTPKLPMDGLFIQTIVAG-MIQLAQSIAKODT---POYNLTIEVSDK----	DPK		430
Db	285	NIROQTTPKPSPNMEYIDPERKQDITVVSPALLRETELLENPKYEIIITQAOWAGADVGLT			344

```

Oy      431 TLCEVOYINVIDINDOIPFPEKSDYCNLTADPMINGSTILITQAOTBDEPFGSSKLYH 490
Db      345 GYATATATVIYDDKXDSPFRTKEF--QATVEGAVG-VYUWLYVEDKODPATGAMRAYT 401
Oy      491 IIKGDEGRHGYVDTDPHTNTGVILIKKLEDEFTAVASNIYFKANPEPLVGYKNYASSF 550
Db      402 IINGNGSGSFELHTNPQINNEGMLSVYKRLDEIGAFHLLIKENEDBPVPDVSYGSSST 461
Oy      551 AKFTLLIVTDVNEAPOFSGHVFQAKVSEDAVIGTKVGNVAKDPEGLD--ISYSLRGDTR 607
Db      462 ATVHIVLVLDVNGPVPFYDPDPMWTKQENISVSGVILVTYNADPDSLQHTIRSVYKDPA 521
Oy      608 GMLKIDHTVGEIJSVAPLDREA---GSPYVQVVAITEVGSSISVSEFHLIMDVNDN 663
Db      522 GMLISINPINGVDTTAVLDRESGPVEHNSVYTALFIALIDSGNPATGTGLMITILEDVNDN 581
Oy      664 PPLIADYVGLFFCHPLSAPSLIPEARLDDOHLFRGPHFPFISGSSSLQND-MEYVKIN 722
Db      582 APVITYP-TVAVBCVDARNLSVYLIGADKDLHPNTPD-FKTEIHKOTVPDVKWKISKIN 638
Oy      723 GTHARLSRHTDPEERAYVYLIRIMDGRPREGIVSLPVYFEGSC 767
Db      639 NTHAVLSLIQ-NINKANTYMLPMTVDSGKPEPTNITDRLNOVCGSC 682

```

RESULT 11	Q9WU10	PRELIMINARY:	PRT: 4589 AA.
ID	Q9WU10		
AC	Q9WU10.		
DT	01-NOV-1999	(TREMBLrel. 12, Created)	
DT	01-NOV-1999	(TREMBLrel. 12, last sequence update)	
DT	01-JUN-2002	(TREMBLrel. 21, last annotation update)	
DE	Protocadherin.		
GN	FAT.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi		
OC	Mammalia; Euteleostomi; Rodentia; Sciurognathi; Muridae; Murinae;		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-SPRAGUE-DAWLEY;		
CC	MEDLINE=99173791; PubMed=10072790;		
RA	Ponassi M., Jacques T.S., Ciani L., French Constant C.;		
RT	"Expression of the rat homolog of the Drosophila fat tumour		
RT	suppressor gene.";		
RL	Mech. Dev. 80:207-212 (1999).		
CC	1- SIMILARITY: CONTRAST 33 CADHERIN DOMAINS:		
DR	EMBL; AF100960; AAD20459.1; --		
DR	HSSP; P01132; IEGF.		
DR	InterPro; IPR000152; Asx_hydroxyl.		
DR	InterPro; IPR002126; Cadherin.		
DR	InterPro; IPR000561; EGF-like.		
DR	InterPro; IPR000742; EGF 2.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	InterPro; IPR001791; Laminin_G.		
DR	Pfam; PF00028; cadherin; 33.		
DR	Pfam; PF00008; EGF; 5.		
DR	Pfam; PF00054; Laminin_G; 1.		
DR	PRINTS; PR00205; CADHERIN.		
DR	SMART; SMO0112; CA; 31.		
DR	SMART; SMO0179; EGF_CA; 1.		
DR	SMART; SMO0201; EGF_Like; 2.		
DR	SMART; SMO0282; Lame; 1.		
DR	PROSITE; PS00010; Asx_HYDROXYL; 1.		
DR	PROSITE; PS00232; CADHERIN_1; 16.		
DR	PROSITE; PS00268; CADHERIN_2; 33.		
DR	PROSITE; PS00022; EGF_1; UNKNOWN_4.		
DR	PROSITE; PS01186; EGF_2; 1.		
DR	PROSITE; PS01187; EGF_CA; 1.		
KW	Calcium-binding; Cell adhesion; EGF-like domain; Glycoprotein		
QO	Hydroxylation; Repeat.		
SEQUENCE	4589 AA; 505987 MW; 8CD0561BDC58677 CRC64;		

Query Match 14.5%; Score 628; DB 11; Length 4589;
 Best Local Similarity 25.3%; Pred. No. 1,3e-31;
 Matches 228; Conservative 135; Mismatches 329; Indels 210; Gaps 32;

26 KFSGPKLPMTFESIVE-----GOEPOIIFQ-----FRANPAVTFELTGETDNIFFVI 72
 2703 RFSEPPF--YTTISBDPIGTEIDLIRVHSGVLYTLVKGNTF-----ESNRDEFVI 2754
 73 ERE-GLLYNRALDRETRSTHNLQVAA---LDANGIIVEGPVPTIEVKDINDNRPFLQ 128
 2755 DROSGRLKLEKSLDHEHTTKWYQFSILARCTLDYEVAAS--IDVSIQVADANDNSPVL 2812
 129 SKTEGSRONSRPKPFLVYNATDLDPATNGQLYYQVYLQPMINWYFOINKTKGA 188
 2813 NPTEAFIVENLPAGSRVIOVRASDLDSGV--NGQWYSL--DQSQDADIIESFALINMETGW 2869
 189 ISLTRREGSOELNAPKPSYMLVSVKDMGQSENSFSDTTSVDIYTE-----NI 238
 2870 ITTLK-----ELDHEERASVQIKVASDHGEKVQ--LSTAIIVDTVTVDNDSPRPTAEI 2923
 239 WKA-----PKVEMVENSSTDPIKITQVRNNDPAQVSLVDEKELPRFPF--SIDQEG 290
 2924 YKGVSEDDPBGVIALISTDADSEINR-----QVSYFITGDMAGQFAVENMQNDW 2977
 291 DIYVTOPLDREKQAVYFAVAKDEYKPLSYLEIHVKYKINDNDPPTC----- 340
 2978 RYVVKKPLDREKQSYLLTATAD--GTFSKARAVEKVLADANDNSPVCERTAYSDAIP 3034
 341 ----- 340
 3035 EDALPGKLWQVATDADIRSNAEITYTLFGSGAEKFKLNDPTGELRTALLDREQAVY 3094
 341 ----- 340
 3095 HLIVKATDGGGRSCQATIVLTLEDVNDNTEPTADPAIATVF--ENTBETPLTRQAT 3151
 366 DRDEENTANSFLNRIYEQTPKLPMDGLFLIQVAGMLAKOSLKQODPRQVYLTIEVS 425
 3152 DADA--GLNRKISLSDSA-----DQOFSINEOSGILQLEKH--LDRELQAVYTLTKAA 3203
 426 D---KDPKTLCFVQIVNDINDIQIPFEKSDYGNLTAEPTNIGSTITLTIOATDADPEF 481
 3204 DQGLPRKLTATGTYVSVLDINDNRPVFEYREYV--ATVSEDIIVIGTEVLQVYASRD--I 3260
 482 TGSGKLIYHIKDSGRLGVDTDPHTNTGVYIIKKPLDFTAAVSNIVKAEPEPLVF 541
 3261 EANAETIYALISGNEHGFSDID---SKTGAIPTIENLDYESSHGUYLVTEATD----- 3310
 542 GVKYNASSFAPKFTLIVTDVNE--APQSOHVQOAVSEDAVIGTVKGVNTAKD---PEGLD 597
 3311 GGPFSLSDVATVNINIDINDNSPVSQDSITYVSEDALEQVITIMDADGPNNSH 3370
 598 ISYS-LKDRTRGMLKIDHVTGEISFVAPLDRSAGSPRVQVAVATEVSGSLSYSEFHLI 656
 3371 ILVSIIEGNGOSPPTIDPVRGEIKVTKPELDBETISGTLTVQADNDGNPRVNTTNTYID 3430
 657 LMDVNDNPRPLAKDYTLGFCHPLSAPGSLIFEATDDQHLFRGPHFTSLGSSGLONW 716
 3431 VSDVNDNAPLFSRDNYSVIIQENKPVGFSVYKLVVTKDSSHNQPPFSFALVSGNDNMF 3490
 717 EVSKINGTNARLSTRHDFE--ERAYVVLIRINDGRRPLEGIYSLVTFSCVEGSGCR 774
 3491 EVND---HGVLLTAATVTKRKVKOHYLLHVAVADNGKPOLSLTHIDI--RVLESIRH 3543
 775 PA 776
 3544 PA 3545

RESULT 12
 090XAX3
 ID 090XAX3 PRELIMINARY; PRT; 4587 AA.

AC 090XAX3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Mouse fat 1 cadherin (Fragment).
 GN MFAT1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
 OC NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=EMBRYO;
 RX MEDLINE=20203461; PubMed=10741417;
 RA Cox B.T.M., Hadjantonakis A.K., Collins J., Magee A.I.;
 RT "Cloning and expression throughout mouse development of mfat1 a
 homolog of the Drosophila tumour suppressor gene fat."
 RL Dev. Dyn. 217:233-240(2000).
 CC -1- SIMILARITY: CONTAINS 33 CADHERIN DOMAINS.
 DR EMBL: AJ250768; CAB65271.1; -
 DR HSSP: P08709; 1BF9.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001881; EGF Ca.
 DR InterPro: IPR001791; Laminin G.
 DR Pfam: PF000028; cadherin; 33.
 DR Pfam: PF00008; EGF; 4.
 DR Pfam: PF00054; laminin G; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA; 31.
 DR SMART: SM00179; EGF CA; 1.
 DR SMART: SM00001; EGF-like; 1.
 DR SMART: SM00282; LamG; 1.
 DR PROSITE: PS00222; CADHERIN 1; 16.
 DR PROSITE: PS0268; CADHERIN 2; 33.
 DR PROSITE: PS00022; EGF 1; UNKNOWN 2.
 DR KAM: Calcium-binding; Cell adhesion; Glycoprotein.
 FT NON TER 4587 4587
 SQ SEQUENCE 4587 AA; 506036 MW; 4D3F23B05127C84 CRC64;

Query Match 14.4%; Score 624; DB 11; Length 4587;
 Best Local Similarity 25.1%; Pred. No. 2.5e-31;
 Matches 227; Conservative 135; Mismatches 329; Indels 212; Gaps 33;

26 KFSGPKLPMTFESIVE-----GOEPOIIFQ-----FRANPAVTFELTGETDNIFFVI 72
 2704 RFSEPPF--YTTISBDPIGTEIDLIRVHSGAVLYLVKGNTF-----ESNRDEFVI 2755
 73 ERE-GLLYNRALDRETRSTHNLQVAA---LDANGIIVEGPVPTIEVKDINDNRPFLQ 128
 2756 DRNGRLKLEKSLDHEHTTKWYQFSILARCTLDYEVAAS--IDVSIQVADANDNSPVL 2813
 129 SKTEGSRONSRPKPFLVYNATDLDPATNGQLYYQVYLQPMINWYFOINKTKGA 188
 2814 SPEAFIVENLPAGSRVIOVRASDLDSGA--NGQWYSL--DQSQDADIIESFALINMETGW 2870
 189 ISLTRREGSOELNAPKPSYMLVSVKDMGQSENSFSDTTSVDIYTE-----NI 238
 2871 ITTLK-----ELDHEERASVQIKVASDHGEKVQ--LSTAIIVDTVTVDNDSPRPTAEI 2924
 239 WKA-----PKVEMVENSSTDPIKITQVRNNDPAQVSLVDEKELPRFPF--SIDQEG 290
 2925 YKGVSEDDPBGVIALISTDADSEINR-----QVSYFITGDMAGQFAVENMQNDW 2978
 291 DIYVTOPLDREKQAVYFAVAKDEYKPLSYLEIHVKYKINDNDPPTC----- 340
 2979 RYVVKKPLDREKQSYLLTATAD--GTFSKARAVEKVLADANDNSPVCERTAYSDTIP 3035
 341 ----- 340
 3036 EDALPGKLWQVATDADIRSNAEITYTLFGSGAEKFKLNDPTGELRTALLDREQAVY 3095
 341 ----- 340

Db 3096 NLVKAIDGGRCQAIVLTLEDVNDNADEFTAEPTTIVF---ENTREGETLTRYQAT 3152
 QY DRDEENTANSEFLNRIYVEOTPKLPMOGLFLIQYAGMLQAKOSLKQODTPQYNLTIEVS 425
 Db 3153 DAD--TGLNRKISYLSVESA-----DQOFISINERSGITOLEKH-LDELOAVTTLTKAV 3204
 QY 426 D----KDFKTLCEVOQINVIDINDOIPFEKSDYGNLTAEADTNIGSTILITIOATDADEPF 481
 Db 3205 DQGLPRLTATGTVVSVLDINDNPPVEFREYEG-ASVSEDIYIGTEVLQVAAASND-1 3261
 QY 482 TGSKLLYHIKIDDSBRLGVDTPHTNTGYVLIKKPLDETAASNVIFKAENPEELVF 541
 Db 3262 ENAAEITVYAIISGNHKGFSID-----SKTGAFIIEISLDYESSHEVYLVEAD----- 3311
 QY 542 GVKYNSAFKFLIYTDVNE-APQSOHVEQAKNSDVAILGKNGVATKD---PEGLD 597
 Db 3312 GGTPLSDVATVNTVINDNPSVSDTYTTSVEDALAEQPVITIMADDDGPNNSH 3371
 QY 598 ISYS-LRCDTRGWLKIDHTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEPHLI 656
 Db 3372 IHSIIIEGNGSGPFTIDPVKGEVKTKEPLDRETISGTLTLVQAADNGNPPRVNTTIVNID 3431
 QY 657 LMDVNDNPPRLAKDYTGLEFCFPLSAPGSLIFEAATDDQHLFRGPHFTPSLGSGLQNDW 716
 Db 3432 VSDVNDNADLFSDHNSVLIQENKPVGFSVLKLVTDKDSHNGPPFFFTIVSGDENAP 3491
 QY 717 EYKINGSTARLSTRHTDPEERA---YVULIRINDGGRPLEGIVSLPVTFCSCVSGSC 773
 Db 3492 EVNQ---HGVLTAAAT-IRKVKDHYFLHVKVADSGKQLSMTHIDI---RVIEBSIH 3543
 QY 774 RPA 776
 Db 3544 PPA 3546

RESULT 13

Q8R508 PRELIMINARY; PRT; 4555 AA.
 AC Q8R508; 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Fact3.
 GN RAT3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;
 RX MEDLINE=21670969; PubMed=11811999;
 RA Mitui K., Nakajima D., Ohara O., Nakayama M.;
 RT "Mammalian fact3: a large protein that contains multiple cadherin and
 RT EGF-like motifs";
 RL Biochem. Biophys. Res. Commun. 290:1260-1266(2002).
 DR EMBL: AB076401; BAB86869.1;
 SQ SEQUENCE 4555 AA; 502077 MW; B47C8C10638EA73B CRC64;

Query Match 14.1%; Score 611; DB 11; Length 4555;
 Best Local Similarity 25.4%; Pred. No. 1.7e-30;

Matches 225; Conservative 144; Mismatches 320; Indels 196; Gaps 33;

QY 30 PLKPMFTSIEGOEBSQIIFQKPNAPVATFELTGSTDNIVYIERE-GLLYNRALDRET 88
 Db 2733 PMSQVRFSTVNGERPE-----NKNENVFIEGTGAIKLDKRLDHEV 2774
 QY 89 RSTNHLQVAA---LDANGIIVGSPVPTIEVKDINDNRTPLQSKKEGVRQNSRPGKPF 145
 Db 2775 SPAPFKVAATPILDKVDIVF--TVDDVVKVLDLNDNKVFETTSYETITIMGMPVGYTL 2832
 QY 146 LVYNAATDLDPATPNGQLYYQIVLIQLPMINNVYFQINKKTAISLTREGSGLNPAKNP 205

Db 2833 AQVRAIDTWGA--NQGVTSLSHSDHLEKWEAFNIDSNTGWIStLK---DLDEHTDP 2886
 QY 206 SYNLVLSVDMGQSGSNSSDTTSVUIITENTAKAPKPEVMEJNSTDPHPIKITOVRAN 285
 Db 2887 TFSFFVAVADLG--EAFSLSSMALVSVKVTIDINDNP-----VFAHEVVRGNVKE 2935
 QY 266 DPG---AQSLVDEKEL-----PRPEFS--IDQEGDIYVTOPLDREKD 304
 Db 2936 DPGEVAVVASTLDKOTSINROVSHITGNGRGRFALGAMQSEKVKVKEPLDREED 2995
 QY 305 AAYFAVADDEYKPLSYLEIHVKVDINDNPTC-----PSPTVEVQOE 351
 Db 2996 IYFLNITASD--GLFVTQM-VEVTVSDVNDNSPVCDOVAYASALPEDIPSNKIILKVA 3052
 QY 352 NER-----LQNS-----IGTLAHRDE----- 369
 Db 3053 KADIGSNGDIRYSLVSGNSDFLDPESGELTKTLALDRERVPVYNLARATDGGGRFC 3112
 QY 370 -----ENTANSFLNRIYVEOTPKLPMO----- 391
 Db 3113 SSTVLALLDNDNPPVSSNHTACVYENTATKALLTRVOAVDPVGINRKVYSLEDS 3172
 QY 392 --GLFLIQYAGMLQAKOSLKQODTPQYNLTIEVSD---KDFKTLCEVOQINVIDINDO 445
 Db 3173 ASGVFSIDSSGVIVL-EQPLDREOQSSYNISVRATDQSGQSLSTVITIVLDINDN 3231
 QY 446 IPFEKSDYGNLTAEADTNIGSTILITIOATDADEPTGSSKLLYHIKIDDSBRLGVD 505
 Db 3232 PVFERRDY-LTVPEPDTSLGTQVLSVFATSKD--LGTNAEITVYLIIRSGNEQCKFRI 3285
 QY 506 PHNNGGVYIKKPLDETAASNVIFKAENPEELVNGVKYNSAFKFLIYTDVNE-AP 564
 Db 3286 -NPKTGISVLEADLEMCCKRFTLVLEAKD-----GTPALSTAATVVIDLTDVNDNP 3338
 QY 565 QFSOHVFOAKVEDVAIGTKVGNVATKD---PEGDISYSY-LRCDTRGWLKIDHTGEI 619
 Db 3339 RFSQDVSAVISEDALMEGDSVILLIAEDVSKPNQ-QIRFSTVGDGRDNEFPAVDPLGLV 3397
 QY 620 FSVAPLDRAGSPYRVQVATEVGGSSLSVSEPHLIIMDVNDNPPRLA-KDYTGLEFCF 678
 Db 3398 KYKKKIDLRERVSGYSLILOAVNSGIPAMSGITTVINDVNDNSVFTPAYTAVIQEN 3457
 QY 679 PLASPGSLIFEAATDDQHLFRGPHFTPSLGSGLQNDWVSKINGHAALSTRHTDPEER 738
 Db 3458 KPVGTSLIQLVVTDRS-FHNQPPPSFSLTSGNEDEEFLDHSGLIRSAVFRHMSPE- 3515
 QY 739 AAVVULIRINDGGRPLEGIVSLPVTFCSCVSGSCFRPACHOTGIP 783
 Db 3516 -VLCTIQADSGKPOO---VSHITYIRVYIESTHKP---TAIP 3552

RESULT 14

Q9V5N8 PRELIMINARY; PRT; 3579 AA.
 AC Q9V5N8; 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE PROTOCADHERIN-like wing protein (FLAMINO protein).
 GN STAN OR CG11895 OR FWI.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=EMERYO;
 RX MEDLINE=20025940; PubMed=10556066;
 RA Chae J.W., Kim M.-J., Geo J.H., Collier S., Gubb D., Charlton J.,
 RA Adler P.N., Park W.J.,

FT CARBOHYD 1441 1441 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1650 1650 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1678 1678 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1747 1747 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1843 1843 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1975 1975 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2016 2016 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2028 2028 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2071 2071 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2088 2088 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2196 2196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2320 2320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2784 2784 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3083 3083 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3228 3228 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 181 181 T -> S (IN REF. 1).
 FT CONFLICT 361 361 O -> L (IN REF. 1).
 FT CONFLICT 395 401 MSGLDS -> NGITVGLP (IN REF. 2).
 FT CONFLICT 1968 1968 Q -> H (IN REF. 1).
 FT CONFLICT 2271 2271 G -> E (IN REF. 1).
 FT CONFLICT 2502 2502 R -> C (IN REF. 2).
 FT CONFLICT 2627 2627 D -> G (IN REF. 2).
 FT CONFLICT 2709 2709 T -> S (IN REF. 2).
 FT CONFLICT 2756 2756 Q -> R (IN REF. 2).
 FT CONFLICT 2901 2901 C -> Y (IN REF. 1).
 FT CONFLICT 3098 3098 L -> P (IN REF. 1).
 FT CONFLICT 3569 3579 ERNIDDETV -> DSEABY (IN REF. 2).
 SO SEQUENCE 3579 AA; 397139 MW; 48801C493031FB19 CMC64;

Query Match 13.9%; Score 601.5; DB 5; Length 3579;
 Best Local Similarity 29.1%; Pred. No. 5e-30;
 Matches 224; Conservative 131; Mismatches 321; Indels 95; Gaps 35;

37 SIYEGEPBQIIQOFANP-----PAYTELGEDN-----IFVIE-REGGLY 79
 472 STIEGATVSGTITVLTATDODICKMAIEIGIAVTDGAGLADDOEMPIFRIDSRGVS 531
 80 YNRALDRESTRNTLQVAALDANGIIVE---GPVITIEKQINDNRPFLQSKYGSVR 136
 532 TSSUDRETSDSHLVLTADLASAQSERRTAASQVAVLDNDNVPQSEETTYVQVP 591
 137 QNSRPG---KPELYNATDIDDPATNGOLYQIVQIPLMINNMVFOINKTGAISLR 193
 592 EIQMGTEEDNTVAHIRATDADQ---GNNAIRYALIG3---NTQSFISDMSGDVSLVK 645
 194 ESGQELNPAKNBPNVYISVKMGQSGSEFSDTTSVDIIVENIMKARK-----PVEM 247
 646 ---PUDYESVRVRLVIRAQDGGSPSR---SNTQLLVAVDANDNAPREFYTSQFOESV 698
 248 VENSTDPHPIKITQVAMNDPGAQ---YSLVDEKELPRFPFSID-OEGDIYVTOPLDRE 302
 699 LENVPGVMIIRVQADSDGNAELTYSISED--DNFPLAVDPRTGWQITKPLDRE 756
 303 KQAYVFAVAKDEYKPLSYPLIEHVKVDINDNPTCSPTVFEVQENRIGNSIGTL 362
 757 QGRFAFQVAVAKGQGVPKSASSSVITVQDVNDNDNAPKPYEAUVAGSDQPGTEVTV 816
 363 TAMDREENTANSFLNRYVEQPKLPMGLPLI--QTAGMLQLAKSLIKODTQVYL 420
 817 TAMPDED---SRLHYETTGTR---GRFITSQNGRGLITTA-QSIDVQKQERKPL 867
 421 TIEVSKDEFT-LCFQVIVINDINDQIFPEKSDYGLTLAEDPTNGITLITQADAE 479
 868 TYAATDSGSRDPAVHINITANNAPIFENAPY-SASVEDAPVGTIVLVASATDS- 925
 480 PTTGSKLIVHIIKQDSERGLVDTP---HTNTGVIIKKPLDFETAASNVIFKXENP 536
 926 --GVVAQITYSINRESINGLG-SPDPFSINPQTGAIVTNAPLDRETTSGYLLTVAKD- 981
 537 EPLVGVKKNASSFAFTLIVTDVNE-AQFSGHVQAVSEVDVAGTGVNTAKDP- 594
 982 -----GNDPSLDTTDEIGVTDVNDNAPAFKSPLYQASILBDALVGTSVIOVAASDPDV 1036

QY 595 GLD-1SYSL--RGDTRGMKIDHTVGEIFSVAPLDRREAGSPYQVAVTEVGGSSLSV 650
 1037 GUNGRKYLISDRDIEDGSFVIDPISGTRTKGDLRESVAVHILAIANDKSSPLSLST 1096
 QY 651 SEPHLLMDVNDNPPPLAVDTGLFCHPLSAP-GLIFE--ATDDQHLFRGPHFTFSI 707
 1097 VEQIRLEEDVNDSPPFASDKITLYV--PENSFVSGVGEIHADDEGNNAVH--YSI 1152
 Db 708 GSGSLQNDME-VSKINGTARIST-RHTDPE--EAYVVLIRINDGCRPL 754
 1153 IGGDSNATSLVTRPGSERRQULTMTEDYESTRKRELVVR--AASPL 1200
 RESULT 15
 ID Q9VW71 PRELIMINARY; PRT; 4643 AA.
 AC Q9VW71
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-JUN-2002 (TrEMBLrel. 13, Last sequence update)
 DE Putative fat-1-like cadherin precursor (CG7749 protein).
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCHI_Taxid:7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Gelinkner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abrial J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borrows K.C., Busam D.A., Butler H., Brockett P., Brothier P.,
 RA Burtis K.C., Bustin M.A., Butler H., Cadieu L.B., Davies P.,
 RA Chery J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster K., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclab J.M.,
 RA Palczolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reiert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svinges R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weissbrock G.M., Weissbrock J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers B.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -!- FUNCTION: MAY BE INVOLVED IN CELL ADHESION.
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE CADHERIN FAMILY.
 CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
 DR EMBL; AE003515; AAP49078.1; -.
 DR HSSP; P15116; INCI.
 DR FlyBase; FBgn0036930; fat2.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00028; cadherin; 31.
 DR Pfam; PF00008; EGF; 5.
 DR Pfam; PF00054; laminin_G; 1.
 DR PRINTS; PRO0205; CADHERIN.
 DR SMART; SM00112; CA_33.
 DR SMART; SM00179; EGF_CA; 1.
 DR SMART; SM00001; EGF-like; 5.
 DR SMART; SM00282; Lamg; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00232; CADHERIN_1; 17.
 DR PROSITE; PS00268; CADHERIN_2; 41.
 DR PROSITE; PS00022; EGF_1; 5.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01187; EGF_CA; 1.
 KW Hypothetical protein; Cell adhesion; Signal; Transmembrane; Repeat;
 KW EGF-like domain; Glycoprotein; Calcium-binding.
 FT SIGNAL 1 35 POTENTIAL.
 FT CHAIN 36 4643 PUTATIVE FAT-LIKE CADHERIN.
 FT DOMAIN 36 1647 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1648 1668 POTENTIAL.
 FT DOMAIN 1669 4643 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 68 180 CADHERIN 1.
 FT DOMAIN 189 288 CADHERIN 2.
 FT DOMAIN 293 397 CADHERIN 3.
 FT DOMAIN 406 504 CADHERIN 4.
 FT DOMAIN 513 610 CADHERIN 5.
 FT DOMAIN 619 713 CADHERIN 6.
 FT DOMAIN 778 874 CADHERIN 7.
 FT DOMAIN 883 977 CADHERIN 8.
 FT DOMAIN 986 1085 CADHERIN 9.
 FT DOMAIN 1094 1190 CADHERIN 10.
 FT DOMAIN 1199 1296 CADHERIN 11.
 FT DOMAIN 1305 1402 CADHERIN 12.
 FT DOMAIN 1413 1503 CADHERIN 13.
 FT DOMAIN 1512 1609 CADHERIN 14.
 FT DOMAIN 1618 1714 CADHERIN 15.
 FT DOMAIN 1723 1812 CADHERIN 16.
 FT DOMAIN 1821 1929 CADHERIN 17.
 FT DOMAIN 1948 2030 CADHERIN 18.
 FT DOMAIN 2024 2137 CADHERIN 19.
 FT DOMAIN 2146 2232 CADHERIN 20.
 FT DOMAIN 2247 2338 CADHERIN 21.
 FT DOMAIN 2347 2465 CADHERIN 22.
 FT DOMAIN 2474 2567 CADHERIN 23.
 FT DOMAIN 2576 2670 CADHERIN 24.
 FT DOMAIN 2679 2779 CADHERIN 25.
 FT DOMAIN 2788 2876 CADHERIN 26.
 FT DOMAIN 2885 2983 CADHERIN 27.
 FT DOMAIN 2992 3083 CADHERIN 28.
 FT DOMAIN 3092 3185 CADHERIN 29.
 FT DOMAIN 3194 3289 CADHERIN 30.
 FT DOMAIN 3298 3394 CADHERIN 31.
 FT DOMAIN 3403 3499 CADHERIN 32.
 FT DOMAIN 3508 3604 CADHERIN 33.
 FT DOMAIN 3618 3671 CADHERIN 34.
 FT DOMAIN 3819 3879 EGF-LIKE 1.
 FT DOMAIN 3881 3919 EGF-LIKE 2.
 FT DOMAIN 3919 4129 EGF-LIKE 3.
 FT DOMAIN 4129 4166 EGF-LIKE 4.
 FT DOMAIN 4243 4279 EGF-LIKE 5.
 FT DISULFID 3823 3835 POTENTIAL.
 FT DISULFID 3830 3867 POTENTIAL.
 FT DISULFID 3869 3878 POTENTIAL.
 FT DISULFID 3885 3896 POTENTIAL.

FT DISULFID 3890 3907 POTENTIAL.
 FT DISULFID 3909 3918 POTENTIAL.
 FT DISULFID 4133 4144 POTENTIAL.
 FT DISULFID 4138 4154 POTENTIAL.
 FT DISULFID 4156 4165 POTENTIAL.
 FT DISULFID 4172 4183 POTENTIAL.
 FT DISULFID 4177 4193 POTENTIAL.
 FT DISULFID 4195 4204 POTENTIAL.
 FT DISULFID 4247 4258 POTENTIAL.
 FT DISULFID 4252 4267 POTENTIAL.
 FT DISULFID 4269 4278 POTENTIAL.
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 843 843 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1106 1106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1198 1198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1312 1312 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1439 1439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1473 1473 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1511 1511 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 4643 AA; 517614 MW; 7EC52BD36ADD4B89 CRC64;

Query Match 13.7% Score 591.5; DB 5; Length 4643;
 Best Local Similarity 24.2%; Pred. No. 3,4e-29;

Matches 210; Conservative 133; Mismatches 312; Indels 193; Gaps 33;

QY 33 PMTFSIVEGO-----EPSQIIFQ-FKANPPAVT-----FELTGETDNI FVI 72
 DB 2764 PVSIEIETDANIPTEKSSVLLKTIESTPFGVTLKLMIGNYTFKSIADQDH-FMI 2822
 QY 73 EREGLLYNRADRETRSTHNOVALDANGIIEGVPVITTEVKDINNRPTELKSYE 132
 DB 2823 SPSEGLILOOTLDREQOESHNLIVAEISTVPFAYADVLLIDVRDENDYEPFDTFYS 2882
 QY 133 GSVRONSRRGKPEFLYVATDLDDEPATNGOLYQYIQLPMINWVYFOINNTGASISLT 192
 DB 2883 ASVAENSEKVISLVKASATDAD--TGPNGDIRYLESPTENIONT--FDIDISGWITLL 2338
 QY 193 REGSOELNPAKPSYVLVSVKMGQSENSFSDTTSVDIIVTEN--IKWAP-KVEVQVE 249
 DB 2939 TSLDRFV---QSEYVFXYIADNGHPKDAKVPVITKIVYDNDAPVFLTPEGISLVE 2994
 QY 250 NSTDEPRPIKITQVRNMDPCAQVSLVDEKLPFR-----PSIDQEGDIYVTOPLDRE 301
 DB 2995 NAL-PSTVLI-MLLIDPD-----IEKQMDFTVSGDKQAOFOJCKSGELFTAKPLDRE 3047
 QY 302 EKDAYVFVAVKDEYKPLSYPLEIHVKVNDINDNDPTCPSPVTFVEVOENERLGSIGT 361
 DB 3048 QLMFPLULSIATD--GR-FTAKANVEIDVYKDINDNDPTCLKPRYHISTNESISIGTTIVE 3104
 QY 362 LTAHDEDEENTANSPLNRYIVEQTPKLPMDGLFIQTVAGMLQLAQSLKKQDTPQYNTLT 421
 DB 3105 VKAIDDPQSKLRFYLSGKAD-----FSIKESGILKVA--SALDRETPPKYKLV 3154
 QY 422 IEVSD-KDRTLCF--VQINVIDINDQIPFEKSDVGNLTLMADNTIGSITLLTQATDAD 478
 DB 3155 AHVQDGDVFOECFSEIITVVDINDNDMPFISAAQY-RVSPVDAQNLTLITVHAMDKD 3213
 QY 479 EPTGSSKILYHIKIDSEGRGLGVDTPHTNTGVYIIRKPLDETAVANSIVFEKAE----- 534
 DB 3214 ---FGVNRQIKSYLMSEHNDFYKIS---KSTQIIRLHSLDRETTISLPRLYTKAEDECGV 3266
 QY 535 -----NPEPLVEGVK----- 544
 DB 3267 PKLHSIATVAVNLTLDNDNPEFSMROYCKILENATHGTEVCKVYATSIDGVNADIHY 3326
 QY 545 -----YNASSF-----AFITLIVTD 559
 DB 3327 FIMSGNEGKFKMDSTTGDLVLANATLDYEMSKFVFLITQAIQDGTPLPLSNNAVYNSITLD 3386

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 00:55:44 ; Search time 44 Seconds
(without alignments)
1817.815 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIQAHLSICLMLYATG.....DKGDNVESAQASEVPLRS 832

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4303	99.6	832	2	S55396
2	3414	79.0	827	2	A53954
3	954	22.1	829	2	I46536
4	715	16.5	913	1	IUCHCR
5	705	16.3	913	1	A47543
6	699.5	16.2	906	1	IUCHCN
7	696	16.1	912	1	IUCHCN
8	694.5	16.1	906	1	IUMSCN
9	688.5	15.9	877	1	IUBOCN
10	684.5	15.8	916	2	C38992
11	677	15.7	783	2	I50116
12	671.5	15.5	826	2	B55363
13	671.5	15.5	886	2	A55363
14	667.5	15.4	847	1	IURHDB
15	667.5	15.4	901	1	IURHDA
16	663.5	15.3	906	1	IUXLC2
17	661	15.3	896	2	I45858
18	657	15.2	712	2	B38992
19	653.5	15.1	809	1	IUBODC
20	653.5	15.1	863	1	IUBODC
21	647.5	15.0	905	1	IUXLC1
22	634	14.7	887	1	IUCHCL
23	630.5	14.6	712	2	IUMSCN
24	630.5	14.6	717	2	I51206
25	609	14.1	780	2	S34438
26	608	14.1	790	2	G02678
27	604	14.0	884	1	IUMSCB
28	597.5	13.8	732	1	IUCHCB
29	591	13.7	882	1	IURHCB

30	585.5	13.6	839	1	IUBODF	desmocollin 1b pre
31	580.5	13.4	895	1	IUXLCP	EP-cadherin precu
32	580	13.4	730	1	IUMSCM	M-cadherin - mouse
33	579.5	13.4	871	2	S47518	cadherin - African
34	576	13.3	761	1	IUBODE	desmocollin 1a - b
35	574.5	13.3	796	2	A53584	OB-cadherin precu
36	574.5	13.3	840	2	I37281	Dec1a precursor -
37	574.5	13.3	894	2	I37282	cadherin-11 - mous
38	572.5	13.2	796	2	I49556	seven-pass transme
39	571.5	13.2	796	2	I48277	desmocollin 1b pre
40	571	13.2	3034	2	T14119	desmocollin 1a pre
41	570	13.2	770	2	B48910	cadherin-15 precu
42	570	13.2	824	2	A48910	cadherin - African
43	565	13.1	814	2	G02878	P-cadherin precu
44	564	13.1	905	2	S43064	
45	560	13.0	822	1	IUMSCP	

ALIGNMENTS

RESULT 1

S55396
LI-cadherin - human
C/Species: Homo sapiens (man)
C/Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 31-Mar-2000
C/Accession: S55396
R/Boettinger, A.; Kreft, B.; Fieger, C.; Dlouhy, B.; Berndorff, D.; Goessner, R.; Taube, A.; Reference number: S55396
Submitted to the EMBL Data Library, December 1994
A/Description: Molecular cloning of human LI-cadherin: evidence for a novel type of cadh
A/Reference number: S55396
A/Accession: S55396
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-832 <BOE>
A/Cross-references: EMBL:X83228; NID:G854174; PIDN:CAA58231.1; PID:G854175
F/455-566/Domain: cadherin repeat homology <CR>

Query Match	Best Local Similarity	Score	Pred. No.	Length	DB 2;	Indels	Gaps
1	MIQAHLSICLMLYATG	99.6%;	4303;	832;			
1	MIQAHLSICLMLYATG	99.5%;	3414;	827;			
61	ELTGETDNI FVIEREGILYYNRALDRSTRSTHNLQVAALDANGI IYEGVPITIEVDIN	22.1;	954;	829;			
61	ELTGETDNI FVIEREGILYYNRALDRSTRSTHNLQVAALDANGI IYEGVPITIEVDIN	16.5;	715;	913;			
61	ELTGETDNI FVIEREGILYYNRALDRSTRSTHNLQVAALDANGI IYEGVPITIEVDIN	16.3;	705;	913;			
121	DNRPTFLQSKYRGSVRQNSRPGKPLYYNATDIDDPATNGQLYYQIVIQLPIMINNVMP	16.2;	699.5;	906;			
121	DNRPTFLQSKYRGSVRQNSRPGKPLYYNATDIDDPATNGQLYYQIVIQLPIMINNVMP	16.1;	696;	912;			
121	DNRPTFLQSKYRGSVRQNSRPGKPLYYNATDIDDPATNGQLYYQIVIQLPIMINNVMP	16.1;	694.5;	906;			
121	DNRPTFLQSKYRGSVRQNSRPGKPLYYNATDIDDPATNGQLYYQIVIQLPIMINNVMP	15.9;	688.5;	877;			
121	DNRPTFLQSKYRGSVRQNSRPGKPLYYNATDIDDPATNGQLYYQIVIQLPIMINNVMP	15.8;	684.5;	916;			
121	DNRPTFLQSKYRGSVRQNSRPGKPLYYNATDIDDPATNGQLYYQIVIQLPIMINNVMP	15.7;	677;	783;			
181	QINNKGAISLTREGSQELNPANPNFYNLVISYKMGCGGSENFSTTSVDITVTENIMK	15.5;	671.5;	826;			
181	QINNKGAISLTREGSQELNPANPNFYNLVISYKMGCGGSENFSTTSVDITVTENIMK	15.5;	671.5;	886;			
181	QINNKGAISLTREGSQELNPANPNFYNLVISYKMGCGGSENFSTTSVDITVTENIMK	15.4;	667.5;	847;			
241	APKPEVMEVNSTDPHPRIKITTOVRWMDPGAQYSLVDEKLP RPFPFSIDQEGDIVYTOPPLDR	15.4;	667.5;	901;			
241	APKPEVMEVNSTDPHPRIKITTOVRWMDPGAQYSLVDEKLP RPFPFSIDQEGDIVYTOPPLDR	15.3;	663.5;	906;			
241	APKPEVMEVNSTDPHPRIKITTOVRWMDPGAQYSLVDEKLP RPFPFSIDQEGDIVYTOPPLDR	15.3;	661;	896;			
301	EEKDAVYFAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTVFEVQENRNLGNSIG	15.2;	657;	712;			
301	EEKDAVYFAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTVFEVQENRNLGNSIG	15.1;	653.5;	809;			
301	EEKDAVYFAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTVFEVQENRNLGNSIG	15.1;	653.5;	863;			
301	EEKDAVYFAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTVFEVQENRNLGNSIG	15.0;	647.5;	905;			
361	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	14.7;	634;	887;			
361	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	14.6;	630.5;	712;			
361	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	14.6;	630.5;	717;			
421	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	14.1;	609;	780;			
421	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	14.1;	608;	790;			
421	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	14.0;	604;	884;			
421	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	13.8;	597.5;	732;			
421	TLTAHRDRENTANSFLNRIYEQPKLPMDGLFLIQTAAQMLQAKOSLKKQDTQYNL	13.7;	591;	882;			

```

Db      421 TIEVSDKPKTLCEVQINVIDINDQPIFEKSDYGNLTIAEDTNGSTILLITQATDDEP 480
      481 FTGSSKILYHIKSDSGRLGVDPDPTNTGYVLIKKPLDEFTAAVNIVKAEKNEPVLV 540
      481 FTGSSKILYHIKSDSGRLGVDPDPTNTGYVLIKKPLDEFTAAVNIVKAEKNEPVLV 540
      541 FGVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTKVGNVTAKDPGLDISY 600
      541 FGVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTKVGNVTAKDPGLDISY 600
      541 FGVCYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTKVGNVTAKDPGLDISY 600
      601 SLRGDTGKWLKIDHVTGEIFSVAPLDRBAGSPYRVQVATEVGSSLSVSEFHLIMDV 660
      601 SLRGDTGKWLKIDHVTGEIFSVAPLDRBAGSPYRVQVATEVGSSLSVSEFHLIMDV 660
      661 NDNPRLAKDYTGIFCHPLSAPGSLFEATDDOHLFRGPHFTPSLGSGLQNDWEVSK 720
      661 NDNPRLAKDYTGIFCHPLSAPGSLFEATDDOHLFRGPHFTPSLGSGLQNDWEVSK 720
      721 INGTARLSTRHTDFEERAYVVLIRINDGRRPLEGIVSLPVTFCSVBSGCFRPAHQOT 780
      721 INGTARLSTRHTDFEERAYVVLIRINDGRRPLEGIVSLPVTFCSVBSGCFRPAHQOT 780
      781 GIPTVGAAGVILTLTLVIGIILAVFIRIKDKGKDNVESQAQSEVKKPLRS 832
      781 GIPTVGAAGVILTLTLVIGIILAVFIRIKDKGKDNVESQAQSEVKKPLRS 832

```

RESULT 2

```

Li-cadherin precursor - rat
N:Alternate names: liver-intestine cadherin
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 31-Mar-2000
C:Accession: A53954
R:Berndorf, D.; Gessner, R.; Kreft, B.; Schroy, N.; Lajous-Detter, A.M.; Loch, N.; Reut
J. Cell Biol. 125, 1353-1369, 1994
A:Title: Liver-intestine cadherin: molecular cloning and characterization of a novel Ca
A:Reference number: A53954; MUID:94266966; PMID:8207063
A:Accession: A53954
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-827 <BBR>
A:Cross-references: GB:X78997; NID:9505562; PIDN:CAAS5631.1; PID:9505563
C:Superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; intestine; liver;
F:1-28/Domain: signal sequence #status predicted <SIG>
F:23-827/Product: Li-cadherin #status predicted <MAT>
F:454-565/Domain: cadherin repeat homology <CR3>

```

Query Match 79.0%; Score 3414; DB 2; Length 827;

Best Local Similarity 78.8%; Pred. No. 1e-206;

Matches 655; Conservative 70; Mismatches 102; Indels 4; Gaps 1;

```

      2 ILQAHLSLCLMLYLATGYGQEGKFGPLKPMFTFSIYEGQEPQIIFQKAPNPAVTFE 61
      1 MWSAQHPLFLCLTLTLGAYGQEGKFGPLKPMFTFSIYEGQEPQIIFQKAPNPAVTFE 60
      62 LTGTDTNIPYEREGLYNRALDRERSTNLTQVALLDANGIIVGEGVPITTEVKIND 121
      61 LTGTDTNIPYEREGLYNRALDRERSTNLTQVALLDANGIIVGEGVPITTEVKIND 120
      122 NRPFLOSKEYGSRONSRPGKFLYVNAATLDDPATPNQOLYQIYIQLPMINNMWYFQ 181
      121 NRPFLOSKEYGSRONSRPGKFLYVNAATLDDPATPNQOLYQIYIQLPMINNMWYFQ 180
      182 INNKTAISLTRREGSOELNPAKNPSYLVISVKMGQSGSENSFSDTTSVDITVTENITWKA 241
      181 INNKTAISLTRREGSOELNPAKNPSYLVISVKMGQSGSENSFSDTTSVDITVTENITWKA 240
      242 PKPEVAMENSTDPHPRIKITQVRANDPQAQYSLVDKEKLPFRPFSIDQSGDIYVQPLDRE 301
      241 PKPEVAMENSTDPHPRIKITQVRANDPQAQYSLVDKEKLPFRPFSIDQSGDIYVQPLDRE 300

```

```

      302 EKDAYVFAVAKDEGKPLSYPLEIHKVVDINDNPTCPSPVTFVEVQENERLGNISIGT 361
      301 EKDAYVFAVAKDEGKPLSYPLEIHKVVDINDNPTCPSPVTFVEVQENERLGNISIGT 360
      362 LTHADDEENTANSLNRYIVBQTPKLPMDGLFLOTITGMLQAKSLKKQOTPOYNLT 421
      361 LTHADDEENTANSLNRYIVBQTPKLPMDGLFLOTITGMLQAKSLKKQOTPOYNLT 420
      422 IEVSDKPKTLCEVQINVIDINDQPIFEKSDYGNLTIAEDTNGSTILLITQATDDEP 481
      421 IEVSDKPKTLCEVQINVIDINDQPIFEKSDYGNLTIAEDTNGSTILLITQATDDEP 480
      482 TGSSKILYHIKSDSGRLGVDPDPTNTGYVLIKKPLDEFTAAVNIVKAEKNEPVLV 541
      481 TGSSKILYHIKSDSGRLGVDPDPTNTGYVLIKKPLDEFTAAVNIVKAEKNEPVLV 540
      542 GVKYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTKVGNVTAKDPGLDISY 601
      541 GVKYNASSPAKFLIYTDVNEAPQSQHVFOAKVSEDVAIGTKVGNVTAKDPGLDISY 600
      602 LRGDTEGKWLKIDHVTGEIFSVAPLDRBAGSPYRVQVATEVGSSLSVSEFHLIMDV 661
      601 LRGDTEGKWLKIDHVTGEIFSVAPLDRBAGSPYRVQVATEVGSSLSVSEFHLIMDV 660
      662 DNBPRLAKDYTGIFCHPLSAPGSLFEATDDOHLFRGPHFTPSLGSGLQNDWEVSK 721
      661 DNBPRLAKDYTGIFCHPLSAPGSLFEATDDOHLFRGPHFTPSLGSGLQNDWEVSK 720
      722 NGTARLSTRHTDFEERAYVVLIRINDGRRPLEGIVSLPVTFCSVBSGCFRPAHQOT 781
      721 NGTARLSTRHTDFEERAYVVLIRINDGRRPLEGIVSLPVTFCSVBSGCFRPAHQOT 780
      782 IPTVGAAGVILTLTLVIGIILAVFIRIKDKGKDNVESQAQSEVKKPLRS 832
      781 IPTVGAAGVILTLTLVIGIILAVFIRIKDKGKDNVESQAQSEVKKPLRS 827

```

RESULT 3

```

Ksp-cadherin - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: I46536
R:Thomson, R.B.; Igarsasi, P.; Biesemeyer, D.; Kim, R.; Abu-Alfa, A.; Soleimani, M.;
J. Biol. Chem. 270, 17594-17601, 1995
A:Title: Isolation and cDNA cloning of Ksp-cadherin, a novel kidney-specific member of
A:Reference number: I46536; MUID:95340560; PMID:7615566
A:Accession: I46536
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-829 <THO>
A:Cross-references: EMBL:U28945; NID:9902885; PIDN:AAC48472.1; PID:9902886

```

Query Match 22.1%; Score 954; DB 2; Length 829;

Best Local Similarity 31.6%; Pred. No. 5.7e-52;

Matches 251; Conservative 142; Mismatches 354; Indels 48; Gaps 22;

```

      62 LTGET-----DNIVIERE-GLLYNRALDRERSTNLTQVALLDANGIIVGEGVPITTEV 116
      57 LSGDSGVAAGPFSVAEGFLVTRALDRERQAEQIQTLEAEQSHVLMGQSYTVHV 116
      117 KQINDRPFLOSKEYGSRONSRPGKFLYVNAATLDDPATPNQOLYQIYIQLPMINN 176
      117 KQINDRPFLOSKEYGSRONSRPGKFLYVNAATLDDPATPNQOLYQIYIQLPMINN 176
      117 KQINDRPFLOSKEYGSRONSRPGKFLYVNAATLDDPATPNQOLYQIYIQLPMINN 176
      177 PVFRLRLEPRALALSPGAGFDHALLEGYQLLVQVQKMGQDA-SGHQATATVETISIVE 235
      237 NIKAKPEVAMENSTDPHPRIKITQVRANDPQAQYSLVDKEKLPFRPFSIDQSGDIYVQ 296
      236 STWTPLEPRVHALLENLKVPRPHILAVQHWGSGDVHYRL---ESQPGPFDVTEGKLYVTG 292

```

RESULT 4

IUCHCR
R-cadherin precursor - chicken
C:Species: Gallus gallus (chicken)
C:date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 16-Jun-2000
C:Accession: JH0424
R:Imazuka, H.; Miyatani, S.; Takeichi, M.
Neuron 7, 69-79, 1991
A:title: R-cadherin: a novel Ca²⁺-dependent cell-cell adhesion molecule expressed in the
A:RefSeq number: JH0424; PMID:91299341; PMID:1712604
A:Accession: JH0424
A:Molecule type: mRNA
A:Residues: 1-913 <INT>
A:Cross-references: GB:D01459; GB:D00849; NID:g222854; PIDN:BAA03356.1; PTD:g222855
A:Experimental source: retina
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
C:superfamily: cadherin; cadherin repeat homology
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; retina; transmembrane
F:1-26/Domain: signal sequence #status predicted <SIG>
F:127-166/Domain: propeptide #status predicted <PRO>
F:167-913/Product: R-cadherin #status predicted <MAR>
F:167-721/Domain: extracellular #status predicted <EXT>
F:169-274/Domain: cadherin repeat homology <CR1>
F:244-249/Region: cadherin binding #status predicted
F:277-389/Domain: cadherin repeat homology <CR2>
F:392-504/Domain: cadherin repeat homology <CR3>
F:507-612/Domain: cadherin repeat homology <CR4>
F:613-721/Domain: cadherin repeat homology <CR5>
F:722-753/Domain: transmembrane #status predicted <TM>
F:754-913/Domain: intracellular #status predicted <INT>
F:870-885/Region: serine-rich
F:280,409,554,629,658,699/binding site: carbohydrate (asn) (covalent) #status predicted

Query Match **16.5% ; Score 715; DB 1; Length 913;**

Best Local Similarity **26.9%; Pred. No. 7,2e-37;**

Matches **203; Conservative 117; Mismatches 323; Indels 106; Gaps 19;**

Dy 123 RPTFLOSRYKSGYRONSRRPGKPLV-----NATDLDDPATNGQLYYQIYI 169
 :|::|||::|
Db 31 KPGFSSEDTAFPSQNIMEGOKLKVFNNCAGKKGVARYETNSLDFFPKRADGTMY--AVH 88
 :|::|||:
Dy 170 QLEPMINNMVFQI-----NNKGALISLTREGSCOR 198
 :|::|||:
Db 89 QVOMASKOILIMTYAMDPOTLGWELVRFLVGEEKLQHGHGHPKRKSGLPVLAQQOSDPT 148
 :|::|||:
Dy 199 LNPAAKPYSNLVSYVKMGGGGSSEMSFSDTTSVDIIIVTENMKAPRPVEVENSTDDHPK 258
 :|::|||:
Db 149 LDEWRHQ-----SAKLRRQRKD-----WTIP-PINPENSRGPFPQO 186
 :|::|||:
Dy 259 ITQVRWN---DFGAQSYSL-VDEKELRPRFPSIDQ-EGDIYVTOPDRREKAAYFYAVA 312
 :|::|||:
Db 187 LVRIISDKDKIEHIRVSITGVADQPMEWEVISIDPVSGRMVTTRMDRERASYNHLRAVA 246
 :|::|||:
Dy 313 KDEYGRLSPLELIHVAKVIDNDNPPTCPSVYTTFEVOENRELGNISIGTLTHADRDEENT 372
 :|::|||:
Db 247 VDMNGKRVENPIDLYIVIIMDNRRPEFINQVANGSVDEGSKFGTYMTVTIANDAADSTY 306
 :|::|||:
Dy 373 ANSFLNVRIYEOTPKLPMDGLFLIQTAGMLQLAKQSLLKQDTPQNYLTIENSDDX---- 428
 :|::|||:
Db 307 ANGMYRIRIYTOQPSSSQNMFTINSSETGDVVYAAGIDREKVCQGYVIAQTAEHGNTLN 366
 :|::|||:
Dy 429 --FKTLCFYOINVINDIQPIFEKSDYGNLTLEDNTIGSTITLCATDADEPFYSSRK 486
 :|::|||:
Db 367 YGLSNATAAIITVDVNNDNPEFFTSTYSG--EVPENRVERVVANLVWDMDQPHSPMN 424
 :|::|||:
Dy 487 ILVHIHKDSSEGSLGVDTDPHTNGYVIKKPLDEFPAVSNIVYKAEPEPLVGEVKN 546
 :|::|||:
Db 425 AIVRIISGPSGSHFTIKRTDPEVTNEGMTVYCAVDYEMNRAMELTVMWSQAELAIGIOXS 484
 :|::|||:
Dy 547 ASSFAKETLIVTDVNEAPOFSOHVFAKVESDVNIGETVGAVTTAKDEGL--DISYSAR 603
 :|::|||:
Db 485 POSTAGTVITSVTDVNERPYPTTNHKLIRLEGVTEGIYLTFPSANDPDREFMOQAVRSYSL 544
 :|::|||:
Dy 604 GTRGMLKIDVHTGELIFSVAPLDREA---GSPYRVQVATEVGGSSLSSVSEFHILMD 659
 :|::|||:
Db 545 SDPANMLINATNGQTTTAVALDRESDYIKNNVYEATFIADNDGIPRASGTGLQIYILD 604
 :|::|||:
Dy 660 VNDNPRELAKDITYGLPFCCHLSAPGSLIFEAT---DDDQULFRGPHFTFSIGS-GSLQ 713
 :|::|||:
Db 605 INDNAELLEPKAQI-----CEKPNLMVINITAADADIDPVV--GP-EVFEELPSVPSAVR 656
 :|::|||:
Dy 714 NDMEVEKINGTHARLSSTRHTDFERBAVVLIRINDGSGPPLIGAVISLPTPCSCGEGSCF 773
 :|::|||:
Db 657 KNWITIIRLNGADVQSLSRIMYTLBAGYDVPIIVDSGNPILNTSIIRKYKCPCDBENGDC 716
 :|::|||:
Dy 774 RPAGHOTGIPTVGMNAVILLTTLVLVIGIILAIVEI 808
 :|::|||:
Db 717 TTIG--AVAAAAGLGTGAITAILICITILTMTVLL 748
 :|::|||:

RESULT 5

A47543
R-cadherin precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 03-May-1994 #sequence_reviseion 26-May-1994 #text_change 22-Jun-1999
C:Accession: A47543; I55501
R:Hultoon, J.C.; Christofori, G.; Chi, W.Y.; Edman, U.; Guest, P.C.; Hanahan, D.; Kelly,
Mol. Endocrinol. 7, 1151-1160, 1993
AtTitle: Molecular cloning of mouse pancreatic islet R-cadherin: differential expression
A:Reference number: A47543; PMID:94067164; PMID:8247017
A:Accession: A47543
A:Molecule type: mRNA
A:Residues: 1-913 <HT>
References: EMBL:X69966; NID:g429111; PID:NCAA49589.1; PID:g429112
R:Matsumami, H.; Miyatani, S.; Inoue, T.; Copeland, N.; Gilbert, D.; Jenkins, N.; Takeic

J. Cell Sci. 106, 401-409, 1993
A>Title: Cell binding specificity of mouse R-cadherin and chromosomal mapping of the gene
A/Reference number: 155501; MUID:94055672; PMID:8270638
A/Accession: 155501
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-913 <RES>
A/Cross-references: GB:014888; NID:9457658; PIDN:BA03605.1; PID:9457659
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F/1-27/Domain: signal sequence #status predicted <PRO>
F/167-166/Domain: propeptide #status predicted <PRO>
F/167-913/Product: R-cadherin #status predicted <MAT>
F/167-721/Domain: extracellular #status predicted <EXT>
F/169-774/Domain: cadherin repeat homology <CR1>
F/244-249/Region: cadherin binding #status predicted
F/277-389/Domain: cadherin repeat homology <CR2>
F/300-304/Domain: calcium binding #status predicted <CR3>
F/392-504/Domain: cadherin repeat homology <CR4>
F/507-612/Domain: cadherin repeat homology <CR5>
F/613-721/Domain: cadherin repeat homology <CR5>
F/722-753/Domain: transmembrane #status predicted <TM>
F/754-913/Domain: intracellular #status predicted <INT>
F/870-885/Region: serine-rich
F/280,409,554,629,658,659/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.3%; Score 705; DB 1; Length 913;
Best Local Similarity 30.1%; Pred. No. 3,1e-36;
Matches 178; Conservative 101; Mismatches 280; Indels 33; Gaps 12;

239 WKAPKPEVENSSTDPHPKITQVRW---NDPGAGVSL--VDEKXIPRPFPSID-QEGDI 232
168 WVIP-PIVVENSGRPPQQLVIRSDKNDIPRISITGVADQPMVEVFNIDSGRM 226
293 YVTPQDLREKDAVYFAVAKDEYKSLPYLEHVVAKDINDNPPCPSPVIVFEQEN 352
227 YVIRPDRERASVHLRAHVNQNGKVENPIDIYIVIDNDNRPFINQVNGSDEG 286
353 ERLGNSIGTLTAHDREENTANSEFLNRYIEQTPKLPMDGLFIQYAGMLQAKOSLKK 412
287 SKRGTYMTVTADADSTANGVRYIYQTPQSPQNMFTINSETGDIIVVAGGLDR 346
413 QDPFQVNLTEVSGD-----FKTLQVQINVIDINDQPIPEKSPYGLTLAEDNTG 466
347 EKVOQYTVIVQATDMEGNLYGLSNVYATITVVDVNDPEFTSTFAG--EVEPKRIE 404
467 STLTQATDADPEFTGSSKILYHIKGDSEGRAGVTDPHNTGYVILKKPLDFETAAV 526
405 TVVANLTVMRDQPHSNMNAVYRIISGDSGHVSRTDVTNEGAVTVKAVDELNRA 464
527 SNIVFAENDEPLVGVKYNASSFAKFTLITVDVNEAPQFSQHVQAKVEDVAIGTKVG 586
465 FMLTVWVNSQAPLAGSIQMSFOSTAGVTITVDVNEAPRPPSNHKLRLREGVAGTALT 524
587 NATKAPREGU---DISLSLGGTRGWLKIDHVTGEISVAPLDREA---GSPRYQVVA 639
525 TFSAVPDRPMQAVRYSKLSDBPAMVHINTNSQITTAAILDRESLYTKNNVYEATFLA 584
640 TEVGGSSLSVSEFHLILMDVNDPRLADYGLFCFPLSAPG--SLIFEATDDQHL 697
585 ADNGIPRPSGTGLIQLIILINDNAPQLRKAQI-----CERGLANAITAADMDP 639
698 FRGPH-FTFSLGSGSLONDEVSKINGTHARLSTRUTFEERAVVILIRINDGRRPLEG 756
640 NIPPVYELPFIPTVAKMWTITRLNGVYALSLRLIYLEAGYDVAVIITDSGNPLSN 699
757 IVSLPVTFCSCVSGSCRRPAGHGTGPTVGMVAVIITLLVIGITLAVPI 808
700 TSVIVKXVCPDENGDCITVG---AVAAAGTGGAIVAILIIVITLIMVLL 748

IJHUCN
Cadherin 2 precursor - human
N/Alternate names: N-cadherin; neuronal cadherin
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 22-Jun-1999
C/Accession: A38870; S11487; J00751; S13799
R/Reid, R.A.
Submitted to the EMBL Data Library, November 1990
A/Reference number: A38870
A/Accession: A38870
A/Molecule type: mRNA
A/Residues: 1-906 <RES>
A/Cross-references: EMBL:X54315; NID:g34998; PIDN:CAA38213.1; PID:g34999
R/Reid, R.A.; Hempeley, J.J.
Nucleic Acids Res. 18, 5896, 1990
A>Title: Human N-cadherin: nucleotide and deduced amino acid sequence.
A/Reference number: S11487; MUID:91016946; PMID:2216790
A/Accession: S11487
A/Molecule type: mRNA
A/Residues: 1-340 'N', 342-698, 'R', 700-704, 'F', 706-906 <RE2>
A/Cross-references: EMBL:X54315
A/Note: this sequence has been revised in reference A38870
R/Walsh, F.S.; Barton, C.H.; Putt, W.; Moore, S.E.; Kelsell, D.; Spurr, N.; Goodfellow, J.
Neurochem. 55, 805-812, 1990
A>Title: N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin
A/Reference number: J00751; MUID:90347462; PMID:2384753
A/Accession: J00751
A/Molecule type: mRNA
A/Residues: 160-194, 'IR', 197-211, 'L', 213-227, 'Q', 229, 'N', 231-235, 'G', 237-248, 'T', 250-356
A/Cross-references: GB:M34064
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion and are thought to
C/Genetics:
A/Genes: GDB:CDH2; NCAD
A/Cross-references: GDB:128185; OMIM:114020
A/Map position: 18q12.1-18q12.1
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
F/1-27/Domain: signal sequence #status predicted <PRO>
F/28-159/Domain: propeptide #status predicted <PRO>
F/160-906/Product: N-cadherin #status predicted <MAT>
F/162-767/Domain: extracellular #status predicted <EXT>
F/237-242/Region: cadherin binding #status predicted
F/270-382/Domain: cadherin repeat homology <CR2>
F/385-497/Domain: cadherin repeat homology <CR3>
F/500-605/Domain: cadherin repeat homology <CR4>
F/606-712/Domain: cadherin repeat homology <CR5>
F/715-746/Domain: transmembrane #status predicted <TM>
F/747-906/Domain: intracellular #status predicted <INT>
F/865-878/Region: serine-rich
F/190,273,325,402,572,622,651,692/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 16.2%; Score 699.5; DB 1; Length 906;
Best Local Similarity 27.4%; Pred. No. 6.7e-36;
Matches 210; Conservative 122; Mismatches 331; Indels 103; Gaps 21;

117 KDINDNPPFTCSVTVFEQENRIGNSIGTLTAHDREENTANSEFLNRYIEQTPKLP 189
47 KDVEGQP-LNVKX-----SNCKGRKQVYSSSEPADKRVDEGQVY--AVRSFPLSS 98
177 VWFQINNTGAIISLTREGSOELNPAKNSVYLVISVKMGQSGSENSFDTSVDTIIV-- 234
99 HAKFLIY-----AQDKETQKQVAVKLSKLT--LTESVYESAVESEIVFP 144
235 -----TENIWKAPKEVENSSTDPHPKITQVRW-----DPGAQ 270
145 RQFSKSHGLQGRKQWVIF-PINLPENSGRPPQQLVIRSDRDKNLSLRYVTGPGA- 202
271 YSLVDEKEXIPRPFPSIDQ-EGDIYVTPQDLREKDAVYFAVAKDEYKSLPYLEHVV 329
203 -----DQPPGTGFIINPISGQLSYTKPLDRQIARFHLRAHAVDINGQVENPIDIVIN 256
330 VKDINDNPPFTCSVTVFEQENRIGNSIGTLTAHDREENTANSEFLNRYIEQTPKLP 189


```

Db      257 VIDANNNDNPEFIHQVWNGTVPEGSKRGTYWTVMTVTHIDADDPPALNGLMRYKIVSQAPESTP 316
QY      390 MDGLFLIQTYAGMLQOLAKOSLKKODTPQYNTLIEVSDX-----FKTLCFVQIVNIDIN 443
Db      317 SPNMFTINNETDDITVVAAGLDREKVOQYTLIIQATMEBGIPPTYGLSNTATAVITVTDVN 376
QY      444 DOIPIF-EKSDVGNLTLMEDNTNIGSTLIIQATDDEPFTSSKILYHIINGDSRGRLGV 502
Db      377 DNPPEFTANTFGEV---PENRVDIIVANLTVDXQPHTPMAVNAVITISGDDPTGRPAI 433
QY      503 DTDPHNTGCVYIIKKPLDPEFAVSNIVFKAKENPEPLFVGVCYNASSFAKFTLATTVDNE 562
Db      434 QTDPSNNGSLVTVVAFIDPETRMFVLVTAHNEQVPLAKGIQHPQGSATATYSVYIVDNE 493
QY      563 APQESQHFQAKVSDVAIGTKVGNVTAKDPEGL--DISYLRGDTRGMLKIDHTGEI 619
Db      494 NPYFAPNPKRIQROEGELHAGTWTLTFTAQDPPRVQOMIRYKTLSDPANMWTIKDIPVNGOI 553
QY      620 FSVAPLIDREA-----GSPYRVQVATVEGSSLSVSSEFHLIMDVNDNPPRLADYTLGF 675
Db      554 TTIAVLDESPVKNKYNINATFTLASDNGIPPSISGCTGLQIYILINDINAPVL----- 606
QY      676 FCHPLSA-----PGLSIFEAETDDOHLFRGPH-FTEPSLGSGLONDMEVSKINGTHAR 727
Db      607 ---PQEAETCENPDNSINITRATDVIDIDPNAGPFAFDLPISVYITKRWITITRLNGDPAQ 663
QY      728 LSTRATDDEBRAYVVLIRINDGRRPPLGCIYSLPTTFSCC-VEGSCFPDAGHQGIPPV- 785
Db      664 LNLKIKFLBAGIYEVPYIITDGSNPPKSNISILRKVCOCOSNGDC-----TDVDRIV 716
QY      786 --GMAVGIILTTLVGTIIIAVVPFRIKKDKGKNVESAQASEVPP 829
Db      717 GAGLGCTGAIIAILICTIITLILVLMFVWMMKRRDKERQAKOLLUD 762

```

```

RESULT 7
L1CCHCN
N-cadherin precursor, neuronal - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1993 #sequence_rev10n 30-Jun-1993 #text_change 22-Jun-1999
C:Accession: A29964
R:Hatta, K.; Nose, A.; Nagafuchi, A.; Takeichi, M.
J:Cell Biol. 106, 873-881, 1988
A:Title: Cloning and expression of cDNA encoding a neural calcium-dependent cell adhesion molecule
A:Reference number: A29964; PMID:88153917; PMID:2831236
A:Accession: A29964
A:Molecule type: mRNA
A:Residues: 1-912 <HAT>
A:Cross-references: GB:K07277; NID:g63649; PIDN:CA02028.1; PID:g63650
C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to be members of a superfamily of calcium-binding, cell adhesion, duplication, glycoprotein, transmembrane proteins.
C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>
F:129-164/Domain: propeptide #status predicted <PRO>
F:165-912/Product: N-cadherin #status predicted <MAT>
F:165-720/Domain: extracellular #status predicted <EXT>
F:167-272/Domain: cadherin repeat homology <CR1>
F:242-247/Region: cadherin binding #status predicted
F:275-387/Domain: cadherin repeat homology <CR2>
F:390-502/Domain: cadherin repeat homology <CR3>
F:503-611/Domain: cadherin repeat homology <CR4>
F:612-720/Domain: cadherin repeat homology <CR5>
F:721-752/Domain: transmembrane #status predicted <TM>
F:753-912/Domain: intracellular #status predicted <INT>
F:869-884/Region: serine-rich
F:278,330,407,578,628,657/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 16.4%; Score 696; DB 1; Length 912;
Best Local Similarity 26.9%; Pred. NO. 1,1e-35;
Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;
126 FLOSKIEGSRQNSRPEKPLVYNATDIDPATNGQLYQIVQLPMINNVYFQIN-- 183

```

```

Db      41  PFDVHSAVSRSHVGGQPLNVRFCODE-----NRKITYGSSSEP 81
      184  -----NKTGAISLTREGSOELNPPANPSYNTLVISYKDMWGQS-----ENSFSDTT 228
Db      82  EDFRVGEDGVVYAEHSPQLSAEPTE-----FVSAARDKETGEHWGMKYKLTPEEPAFGTAS 136
Qy      229  SVDIITVTENI-----WKAPEFVEVENSTDPHPIKTIQVRNNDPCA 269
Db      137  EKQCKKIDIEDIIFPMQOYDSSHLRQKRDWVIP-PINLPENSRGPFPOELVIR----- 189
Qy      270  QYSLVDEKELRPFPSI-----DQ-----EGDIYVQPLDRREKAYVYAAK 313
Db      190  ----SDRKXSLSRYSVYGPAGDOPGTIFINPISQGLSVYKPLDRQOLASFHLRAHAV 245
Qy      314  DEYKPELSYPIEIHVKVNDINPPTCSPTVFPEVOENRLEGNISGTLTADRDREBNTA 373
Db      246  DVNGQVENPIDIVYINVTIDMNDNRBEFLHQWNGVTPBGSFKGYTVMVTALADADPNQA 305
Qy      374  NSFNLRYRIVEQTPKLPMDGLFLIQTAYAGMLQOLAKOSLKODTPQYNLTIEVSDKD----- 428
Db      306  NGMLRYRLISQAPSPSPMFTINNETGDIITVAAGLDRBKQOQTLLIITQATDMGNGPTY 365
Qy      429  -FKTLCFYQVINVIDNDQIPIF-EKSDYGNLTLAEDNIGSTILTITQATDADEPTGSSK 486
Db      366  GLSNATATAVITVTVDVNDNPEFTAMTFYGEV---PENRVADYIVANLTVTDXKQPHTPAMN 422
Qy      487  ILVHIHKSDSGBRGVDPDPTNNGYVILIKKPLDEPTAAVSNIVKAEANPELVAGVCKN 546
Db      423  ARYQQTGGDPTGQFTILTDPNNSDGLVTVPKPIDEFETRMFVLTVAAENOVPLANGIHP 482
Qy      547  ASSFAKFLIITVDVNEAPQSOHVFOAKVSEDAVIGTVGNVTAKDP-----EGLDISYSL 602
Db      483  PQSTRVTSYITVIDNESPYFVNPFLVNOBESGLAGSNLTTFTIAPDPORIMQOTSILRSK 542
Qy      603  RGDTRGMKIDHYTGEIERSVAPLDRFA-----GSPRYOVVAITEVGGSSLSASVEFHILIM 658
Db      543  LSDPANWLKIDPVNGQIITTAVALDRESIYQNNMYNAFLASDNQIPMSGCTGIQIYIL 602
Qy      659  DVNDNPPRL-AKDYTGLEFCHPLSAPGSLFEATDDOGLFRGPHFTSLASG-SLOND 715
Db      603  DINDAPQVNPREAT---TCETLQ-PNAINITAVDPIDPNAGP-FAFELPDSPSIKRN 657
Qy      716  MEVSKINTHARLSRTRHDPFEERAYVULIRINDGRPRPLEGIVSLPTVFCSC-VEGSCFR 774
Db      658  WTVIRISGDHQLSRIRFLBAGIYDPIVITDSGNPFASTSVLKVVCOCODINDGC-- 715
Qy      775  PAGHQGTGPEV---GMAVGILTLTLVGIILAVVPIRIKDKGDNVESAOASQSEVRK 829
Db      716  ----TDVDRIVVAGLGTGATIAIILLCIILITLVLEFVVMVKRRDKERQAKQLIIDP 768

```

RESULT 8
 IJMSGN
 N-cadherin precursor, neuronal - mouse
 C.Species: Mus musculus (house mouse)
 C.Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
 C.Accession: A32759; A46163
 R.Miyatani, S.; Shimamura, K.; Hatata, M.; Nagafuchi, A.; Nose, A.; Matsunaga, M.; Hatata
 Science 245, 631-635, 1989
 A.Title: Neural cadherin: role in selective cell-cell adhesion.
 A.Reference number: A32759, MUID:69346746; PMID:2762814
 A.Accession: A32759
 A.Molecule type: mRNA
 A.Residues: 1-906 <M>
 A.Cross-references: GB:M31131; NID:G192327; PIDN:AAA37353.1; PID:G309125
 R.Miyatani, S.; Copeland, N.G.; Gilbert, D.U.; Jenkins, N.A.; Takeichi, M.
 Proc. Natl. Acad. Sci. U.S.A. 89, 8443-8447, 1992
 A.Title: Genomic structure and chromosomal mapping of the mouse N-cadherin gene.
 A.Reference number: A46163, MUID:92409532; PMID:11528649
 A.Accession: A46163
 A.Status: preliminary
 A.Molecule type: DNA

A:Residues: 839-906 <M12>
 A:Cross-references: GB:545011; NID:g256010; PID:AA823356.1; PID:g256011
 A>Note: Sequence extracted from NCBI backbone (NCBIN:1137519, NCBI:P:113760)
 C:Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought to
 C:Superfamily: cadherin; cadherin repeat homology
 C:Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pr
 F:1-27/Domain: signal sequence #status predicted <SIG>
 F:28-159/Domain: propeptide #status predicted <PRO>
 F:160-906/Product: N-cadherin #status predicted <MAT>
 F:160-906/Domain: extracellular #status predicted <EXT>
 F:162-267/Domain: cadherin repeat homology <CR1>
 F:237-242/Region: cadherin binding #status predicted
 F:270-382/Domain: cadherin repeat homology <CR2>
 F:385-497/Domain: cadherin repeat homology <CR3>
 F:500-605/Domain: cadherin repeat homology <CR4>
 F:606-714/Domain: cadherin repeat homology <CR5>
 F:715-746/Domain: transmembrane #status predicted <TM>
 F:747-906/Domain: intracellular #status predicted <INT>
 F:865-878/Region: serine-rich
 F:190,273,325,402,572,651,692/Binding site: carbohydrate (asn) (covalent) #status predic

Query Match 16.1%; Score 694.5; DB 1; Length 906;
 Best Local Similarity 27.6%; Pred. No. 1.4e-35;
 Matches 211; Conservative 122; Mismatches 332; Indels 99; Gaps 23;

117 KDINDNPTFLQSKYSGSVNRGKPFLLVYNAITDLPATPGOLYQIVQLPMINN 176
 47 KDVEGQP-LINVFSSNCRK-----KRVQYESSPADFKVDEGITY--AVRSPFLTAE 98
 177 VMVFOI---NNKTG-----AISTREGSQELNPAKNP-SYNIIVISYKDM---GGQSENSE 224
 99 QAKFLIYAQDKETQEKQVAVNLSREPLTEBPKEHEIEIYFPQLAGSALROK 158
 225 SDTTSVDIIVENIWKAKPYEMVENSIDPHIKITQVRN-----DPAQOYS 272
 159 RD-----WVLP-PINLPENSNGPPOELVIRSDRDKNLSRYSVTGPGA--- 202
 273 LVDKEKLPRFPFSIDQ-BGDIYVTOPLDREKDAVYFAVAKDEYKPLSLPIELHYK 331
 203 ---DQPTGIFIIINPISGQSVTKPLDRELIAFHRLAAHADVINGQVNPIDIVNI 258
 332 DINDNPTCPSPVTVFEVQENRIGNSIGTLTAHDREENTANSFLNRIYEQPKLPMD 391
 259 DMNDNRPEFLHQVWNGSVPEBSKPGTYMTVTAIDADDPNALNGMLRILISQAPSPSP 318
 392 GLFLIQTAYAGMLQAKOSLKKQDTPQYNLTIEVSKD-----FKTLCFVOINVIDNQ 445
 319 NMFTINNETGDIITVAAGLDREKVOQYTLITQATDMENPTYGLSNTATAVITVTDVND 378
 446 IPIF-EKSDYGNLTIAEDTNGSTLTITQATDADPEFTGSSKILYHIKGSBGLGVT 504
 379 PPEFTAMTFYGEV---PENRVDVIYANLTVDKQDPHTPANNAIYRISGGDPGTFALIT 435
 505 DPHNTGTVIITKKPLDEFTAAVSNIVFAENPEPELVFGVKNASFAKFTLITVDNEAP 564
 436 DPNSNDGLVTVVKKPIDFTNRMFVLTVAAENQVPLAKIGHPPOSTAVSVTVDIVNEP 495
 565 QFSQHVFOAKVSEDAVIGTGVNTAKDPEGL---DISYSLRGDTRGWLKIDHVTGEIS 621
 496 YFAENPKIIRQEBGLHAGTMTLTLAODPRYMOQNIIRYTKLSDPANMLKIDPVNGQIT 555
 622 VAPLDREA---GSPYRVQVATEVGVSSLSVSFHLIMDVNDNPRRLAKDVTGLFCF 677
 556 IAVLDRESPPVQNNIYNATFLASDNGIPMGSGTGLTQIYLLDINDNNAQV----- 606
 678 HPLSA-----PGSLIFEATDDOHLFRGPH-FTFSIGSGSLQDMVEVSKINGTHALS 729
 607 -PQAEETETPEPNSINIALDDYDIDPNAGPAPFLPLSPYTIKRWITINLNGDPAQIN 665
 730 TRHTDFEERAYVTLIRINDGAPPLIEGIVSLPVTFQSC-VEGSGCFRPAHQGTGIPV--- 785
 666 LKIFLEAGIYEVPIITIDSGNPKSNIISIRVKVQCQDSNGDC-----TDVDRIVGA 718

QY 786 GMAVGILITLTVIGIILAVFIRIKDKGKDNVESQAQSEVPK 829
 DB 719 GLGTGALINALICITILLIIVLMFVVMKRRDRERQAKULLIDP 762

RESULT 9 IJBON

N-cadherin precursor - bovine (fragment)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 30-Jun-1993 #sequence_revstion 30-Jun-1993 #text_change 22-Jun-1999
 C:Accession: S11693
 F:1-130/Domain: propeptide (fragment) #status predicted <MAT>
 F:131-877/Product: N-cadherin #status predicted <EXT>
 F:131-877/Domain: extracellular #status predicted <EXT>
 F:133-238/Domain: cadherin repeat homology <CR1>
 F:208-313/Region: cadherin binding #status predicted
 F:241-353/Domain: cadherin repeat homology <CR2>
 F:356-468/Domain: cadherin repeat homology <CR3>
 F:471-576/Domain: cadherin repeat homology <CR4>
 F:577-685/Domain: cadherin repeat homology <CR5>
 F:686-717/Domain: transmembrane #status predicted <TM>
 F:718-877/Domain: intracellular #status predicted <INT>
 F:836-849/Region: serine-rich
 F:161,244,296,373,543,593,622,663/Binding site: carbohydrate (asn) (covalent) #status p

Query Match 15.9%; Score 688.5; DB 1; Length 877;
 Best Local Similarity 26.7%; Pred. No. 3.1e-35;
 Matches 203; Conservative 126; Mismatches 343; Indels 87; Gaps 19;

126 FLQSKYSGSVNRGKPFLLVYNAITDLPATPGOLYQIVQLP-----MINN 176
 7 FPEVDVSAVLSRDLVGQPLNKFNSNC-----NKRKQYESSPADFKVDEGKVYA 60
 177 VMVFOINNKGAISLREGSQELNPAKNPEYNIIVISYKDMGGQSENSEFSDTSDIIV 234
 61 VRSPFLSSSEHSKFLIYAQDK-----TQEKQVAVNLSKPLAPBOSVKSREIBEVFP 115
 235 -----TENIWKAKPYEMVENSIDPHIKITQVRN-----DPAQ 270
 116 RQYTKNGVYQKQKRWVLP-PINLPENSNGPPOELVIRSDRDKNLSRYSVTGPGA- 173
 271 YSLVDKEKLPRFPFSIDQ-BGDIYVTOPLDREKDAVYFAVAKDEYKPLSLPIELHYK 329
 174 ---DQPTGIFIIINPISGQSVTKPLDRELIAFHRLAAHADVINGQVNPIDIVNI 227
 330 VKDINDNPTCPSPVTVFEVQENRIGNSIGTLTAHDREENTANSFLNRIYEQPKLP 389
 228 VIMNDNRPEFLHQVWNGVPEBSKPGTYMTVTAIDADDPNALNGMLRILISQABSTP 287
 390 MDGLFLIQTAYAGMLQAKOSLKKQDTPQYNLTIEVSKD-----FKTLCFVOINVIDN 443
 288 SPNNFTINNETGDIITVAAGLDREKVOQYTLITQATDMENPTYGLSNTATAVITVDN 347
 444 DQIPIF-EKSDYGNLTIAEDTNGSTLTITQATDADPEFTGSSKILYHIKGSBGLGV 502
 348 DNPEFTAMTFYGEV---PENRVDVIYANLTVDKQDPHTPANNAIYRISGGDPAGFAI 404
 503 DTPHNTGTVIITKKPLDEFTAAVSNIVFAENPEPELVFGVKNASFAKFTLITVDNE 562
 405 QTDPSNDGLVTVVKKPIDFTNRMFVLTVAAENQVPLAKIGHPPOSTAVSVTVDIVNE 464
 563 APQSGVFOAKVSEDAVIGTGVNTAKDPEGL---DISYSLRGDTRGWLKIDHVTGEI 619

```

Db      465 NPYFAPPKIRI08EGHAGTGLTFTTAQDDPRYMOQNIIRYTKLSDPANMLKIDSVNGOI 524
Qy      620 FSVAPLDRRA-----GSPYRQOVAVTEVGGSSLSVSEFHLIMDVNDNPPRLADYGLF 675
Db      525 TTIAVLDRRESNNVKAANYMTFLASDNGIPMSGTGLQITLLINDNAQOVLPQEMAFI 583
Qy      676 FCHPLSAPGSLIFEATDDOHLFRGPH-FTFSLGSGSLQNDWEVSKINGTHARLSTRTD 734
Db      584 -CE-TPDPSNINITATLYDIDPNAGPFAFLPLSPVTKIKRWITIRLNGDPAQNLKIKF 641
Qy      735 FEERAYVVLIRINDGRRPLEGIYSLEPYTCSC-VBESSCRPAHQIGIPV---GNVAG 790
Db      642 LEAGIYEVPIITITSGNPPKSNISILRYKVCQCCSDNDG-----TDVDRIVAGLGIG 694
Qy      791 ILTTLTVIGIILAVFIRIKDQKDNVESQAQSEVYP 829
Db      695 AIIAIIICITLILVIMFVVMKRRKROAKQLIDP 733

```

RESULT 10

```

C38892
Cadenherin 4 precursor - human
N/Alternate names: R-cadherin
C/Species: Homo sapiens (man)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 20-Aug-1999
C/Accession: C38892
R/Suzuki, S.; Sano, K.; Tanihara, H.
Cell Regul. 2, 261-270, 1991
A/Title: Diversity of the cadherin family: evidence for eight new cadherins in nervous t
A/Reference number: 524305; MUID:91283540; PMID:2059658
A/Accession: C38892
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-916 <SUZ>
A/Cross-references: GB:U34059, NID:9506409, PIDN:AAA35627.1, PID:9506410
C/Genetic:
A/Gene: GDB:622850
A/Cross-references: 16q24.1-16pter
A/Map position: 16q24.1-16pter
C/Superfamily: cadherin repeat homology
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-166/Domain: propeptide #status predicted <PRO>
F:167-913/Product: R-cadherin #status predicted <MNT>
F:167-721/Domain: extracellular #status predicted <EXT>
F:169-724/Domain: cadherin repeat homology <CR1>
F:244-349/Region: cadherin binding #status predicted
F:277-389/Domain: cadherin repeat homology <CR2>
F:300-504/Domain: calcium binding #status predicted <CR3>
F:392-504/Domain: cadherin repeat homology <CR3>
F:507-612/Domain: cadherin repeat homology <CR4>
F:613-721/Domain: cadherin repeat homology <CR5>
F:722-753/Domain: transmembrane #status predicted <TMN>
F:754-913/Domain: intracellular #status predicted <INT>
F:870-885/Region: serine-rich
F:283,412,557,632,661,702/Binding site: carboxydrate (Asn) (covalent) #status predicted

```

```

Query Match      15.8%; Score 684.5; DB 2; Length 916;
Best Local Similarity 29.3%; Pred. No. 6e-35;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

Qy      239 WKAPKPEVWENSTDPHPIKITQVRM---NDPGAQVSL--VDKEKLPKRPFSID-QEGDI 292
Db      171 WVIP-PNVNPNRGRPPQOLVIRSDKNDIPRISITGVGADQPEVFSINSNGRM 229
Qy      293 YVTOPLDREKDAYVFAVAKDEYKPLSYPLHIVKVDINDNPTCPSPVTVFEVOEN 352
Db      230 YVTRPMDREHHAHYHLRAHVAVDNMGKVENPIDLIYIVIDMNDNHPFINOVANCSDRG 289
Qy      353 ERLGNSIGTLTAHDREENTANSFLNRYIEQTPKPLPMDGLFIQTYAGMLQAKOSLKK 412
Db      290 SKPGIYVMTTANDADSTYANGVVRVYIVTQTPQSPQNMFTINSETGDIVVVAAGWR 349

```

```

Qy      413 QTPPYNLTIKESDKD-----FKTLCFQVQINVIDINDQIPIPEKSDYGNLTAEPTNIG 466
Db      350 EKVOQTAVIQAQDMEGNLNGVSNATATLITVDVNDNSETASTAGPAG--EVPENSVE 407
Qy      467 STLTITQATDADDEPFGSSKILYHIHKGDEBGLVDTDPHNTGYIYIKKPLDFTAAV 526
Db      408 TVANLVMDRDOPHSPNNAAVYRIISGDSGHFVRTPDVTNEGWTYKAVDYEINRA 467
Qy      527 SNIVFAENEPILVFGVKNVASSFAKFTLLIVDVNEAPQSSQVPOAKVSEDAVIGTKV 566
Db      468 FMITVWSNOAPLASGLOQSPSTAGVYISIMDINAPYFSPSHKILRLEGVPPGTVLT 527
Qy      587 NTAQDPEGI---DISYLRGDTFRGLKIDHVGIEFSVAPLDRRA-----GSPYRQOVA 639
Db      528 TFSVAVDPDRMQQAVAYNSKLSDSPASLHNATNGQITTVAVDRELSYTGNNYEATFLA 587
Qy      640 TEVGGSSLSVSEFHLIMDVNDNPPRLADYGLFFCHPLSAPGSLIFEATDDOHLFR 699
Db      588 ADNGIIPASGTGTLQITVLLIDINDNAPELPRKQAICERPLNA---INTPAADADYHPNI 644
Qy      700 GPH-FTFSLGSGSLQNDWEVSKINGTHARLSTRTDFFERAYVVLIRINDGRRPLEGIY 758
Db      645 GPYVFEPLPVPAARVKNWITIRLNGDYAQLSLIYLEAGMTDVPITVDSGNPLSNIS 704
Qy      759 SLPTVTCSCVSGSCFRPAGHQIGIPYGNVAVGILLTTLTVIGIILAV--FIRIKDKG 816
Db      705 IIRKVKCPDDNDGCTTIG---AVAAAGTGAIVAILICITLILVIMLFVVMKRRK 761
Qy      817 D 817
Db      762 E 762

```

RESULT 11

```

I50116
N-cadherin precursor - zebra fish
C/Species: Brachydanio rerio (zebra fish)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jan-2000
C/Accession: I50116
R/Bitzur, S.; Kam, Z.; Geiger, B.
Dev. Dyn. 201, 121-136, 1994
A/Title: Structure and distribution of N-cadherin in developing zebrafish embryos: morp
A/Reference number: I50116; MUID:95178741; PMID:7873785
A/Accession: I50116
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-783 <BIT>
A/Cross-references: EMBL:X67648; NID:9623209; PIDN:CAA47890.1; PID:9623210
C/Superfamily: cadherin repeat homology
F:146-258/Domain: cadherin repeat homology <CR>

```

```

Query Match      15.7%; Score 677; DB 2; Length 783;
Best Local Similarity 30.1%; Pred. No. 1.4e-34;
Matches 187; Conservative 98; Mismatches 288; Indels 48; Gaps 15;

Qy      239 WKAPKPEVWENSTDPHPIKITQVRM---IKITQVRKNDGAQVSLV--DKEKLPKRPFSID-QEGDI 292
Db      37 WVIP-PNVNPNRGRPPQOLVIRSDKNDIPRISITGVGADQPEVFSINSNGRM 229
Qy      293 YVTOPLDREKDAYVFAVAKDEYKPLSYPLHIVKVDINDNPTCPSPVTVFEVOEN 352
Db      96 SYTKPLDRKHIRNFRHLRAHVAVDINGQNMENPIDLIINVIDMNDNHPFEHTQINGTVDBG 155
Qy      353 ERLGNSIGTLTAHDREENTANSFLNRYIEQTPKPLPMDGLFIQTYAGMLQAKOSLKK 412
Db      156 AKRGTFVMTVTSQDKDDPNTANGMLRYKITLSQTPSPSSNMFTINNKTKIITVAAGLDR 215
Qy      413 QTPPYNLTIKESDKD-----FKTLCFQVQINVIDINDQIPIPEKSDYGNLTAEPTNIG 466
Db      216 EKVPQYTLITQATDEGNFTYGLSNATAVIRLLDVNDNAPFTRETFEGVEPNNRVNYI 275
Qy      467 STLTITQATDADDEPFGSSKILYHIHKGDEBGLVDTDPHNTGYIYIKKPLDFTAAV 526

```

```

Db      276 VTMLTV--TDKDEGTAPMAAVYRIISGDPGRSIPITDPTVMEGLTVVVKPVDFEENRS 333
      527 SNIVFAKNEPEPLVFGKYNASSFAKTLTYTVNEAPOSQAHVFOAKVEDALIGKVG 566
      334 FMLTVADVNEPVLASGTHRTROSTATVSIKVIDNEBPNFDPNKOIKLEBGLPQWMLT 393
      587 NVTA-KOPEGL---DISYSLRGDTRGMLKIDHVTGEISVAPLUREA---GSPYRQVW 638
      394 TFTAHDPPDRYMOQTISYSKLYDPANLLEIDPNNGRSTIAVLDRSPYKANNLYATPM 453
      639 ATEVGSLSLSSVSEFHLILMDVNDNPRLAKDYTLGFECPLSA-----PGSLIPEAT 691
      454 ASDNGVPRASGTGLQIYLLDINDNARV-----PQGEAVGCEPPEPAINITAV 503
      692 DDDOHLFRGPHFTSLSSG--GSLONDEVSKINGTHARLSRTHDPEERAVVLIRINDG 749
      504 DGDINPNVAGP-YAFELNRPBIDIRNWTLTRISGDHQLSKIKSYLSGIELEFISYDS 562
      750 GRPLEGIVSLPVTFGSC-VEGSCFRPAGHOTGIPTVGMVGLITLLVIGIILAVFI 808
      563 GNLPMSNTTYLRIRVCCQDHHGDCV---DMERIMAGLGTGAILIILCIILVLVM 618
      809 RIKDKGKDNVESAQASEVKP 829
      619 FVMMKKRRDKERQAKOLLIDP 639

```

RESULT 12

```

B55363
desmocolin, type 4, short form precursor - human
C/Species: Homo sapiens (man)
C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-May-1997
C/Accession: B55363
R;Kawamura, K.; Watanabe, K.; Suzuki, T.; Yamakawa, T.; Kamiyama, T.; Nakagawa, H.; Tsui
J. Biol. Chem. 269, 26295-26302, 1994
A>Title: cDNA cloning and expression of a novel human desmocolin.
A/Reference number: A55363; MUID:95014464; PMID:7929347
A/Accession: B55363
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-826 <KAW>
A/Cross-references: GB:D17427
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: alternative splicing; transmembrane protein
F;138-243/Domain: cadherin repeat homology <CRI>
F;246-355/Domain: cadherin repeat homology <CR2>

```

```

Query Match      15.5%; Score 671.5; DB 2; Length 826;
Best Local Similarity 30.8%; Pred. No. 3.3e-34;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

```

```

      239 WKAPKEVEMVNSTDPHPKIKITQVRWMDPGAQYSL-----VDKEKLPFPFSDIOEG 290
      137 W-APICSMQENSIGPPLFLQVE-SDAQNVTVFYSISGRGVDKPEPLNFYIERD-TG 193
      291 DIYVTOPLDREKDAY--VFYAVAKDEYKGLSPLLEIHVKVQDINDNPPTCSPYVFE 348
      194 NLFCRTPVREBYVFDLIAVASTADGSADL--PLPLRIVDEDNHVFTEATYNFE 251
      349 VOENBERGNSIGTLTAHDBEENTANSFLNRYIVEOPKLPMDGLFLIQTAGMLQALAKQ 408
      252 VLSSSRGTTVGVGVCATDRDEPDMTRKLSIILOQTPRSP--GLFSVHPTGVIITVSH 309
      409 SLKQDTPQYNLTIEVSDKPK-----TLCFVQVINVDINDQIPFEKSDYGNLTALAE 461
      310 YLDREVVDKXSLIMKVQMDGQFGILGSTSCI--ITVDSNDNAPFRQNAVE--AFVE 365
      462 DTNIGSTILTQATDADEPFTGSSKILYHIKDSBGLGVDPDPTHTNTGVIYIKPLDF 521
      366 ENAFNVEILRIPIEDKDLINTANRWVNTILKGNENGHFKISTDKETNEGVLSVVKPLNY 425
      522 ETAAVSNIVFAKNEPEPLVFGV-KYNASSFAKTLTYTVNEAPOSQAHVFOAKVEDVA 580

```

```

      426 EENRQVNLIEGVNNEAPFARDIPRYALNRAVTVHVRDIDDEGEPECTPAQAVYRIKENLA 485
      581 ICFKVGNTVAKOPE---GDISYSLRGDTRGMLKIDHVTGEISVAPLUREAGSP-----Y 633
      486 VSGIKNGYKAYDENENNGRLRKJLHDPGWTITDEISSGITSKILBREVTPOKELY 545
      634 RVGVNVADEVGSSLSVSEFHLILMDVNDNPRLAKDYTLGFECPLSAAGSLIPEATD 693
      546 NITVALAIDDDRSCTGLAVN--IEDVNDNPPEILQEV--VVICRKMGVTDIL--AVDP 599
      694 DQHLFRGPHFTSLSSGS--LONDEVSKINGTHARLS-TRHTDEERAVVLIRINDG 750
      600 DEPVHGAP-FYFSLPNTSPEISRLMSLTLYVNDPAARLSQKNAQFOE--YTIPIYKXD-- 654
      751 RPLLEGIVSLPVTFGSCVSGSCFRPAGHOTGIPTVGMVGLITLLVIGIILAVFI 809
      655 RAGQAAATKLLRVNLCCHPTQCRATRSSTIGVLGKMAIALLILALFSLVLTIVCGV 714
      810 IKDKGK 816
      715 FGATKGR 721

```

RESULT 13

```

A55363
desmocolin, type 4, long form precursor - human
C/Species: Homo sapiens (man)
C/Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Jun-2000
C/Accession: A55363
R;Kawamura, K.; Watanabe, K.; Suzuki, T.; Yamakawa, T.; Kamiyama, T.; Nakagawa, H.; Tsui
J. Biol. Chem. 269, 26295-26302, 1994
A>Title: cDNA cloning and expression of a novel human desmocolin.
A/Reference number: A55363; MUID:95014464; PMID:7929347
A/Accession: A55363
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-896 <KAW>
A/Cross-references: GB:D17427; NID:9639672; PIDN:AAA04249.1; PID:9685219
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: alternative splicing; transmembrane protein
F;138-243/Domain: cadherin repeat homology <CRI>
F;246-355/Domain: cadherin repeat homology <CR2>

```

```

Query Match      15.5%; Score 671.5; DB 2; Length 896;
Best Local Similarity 30.8%; Pred. No. 3.8e-34;
Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

```

```

      239 WKAPKEVEMVNSTDPHPKIKITQVRWMDPGAQYSL-----VDKEKLPFPFSDIOEG 290
      137 W-APICSMQENSIGPPLFLQVE-SDAQNVTVFYSISGRGVDKPEPLNFYIERD-TG 193
      291 DIYVTOPLDREKDAY--VFYAVAKDEYKGLSPLLEIHVKVQDINDNPPTCSPYVFE 348
      194 NLFCRTPVREBYVFDLIAVASTADGSADL--PLPLRIVDEDNHVFTEATYNFE 251
      349 VOENBERGNSIGTLTAHDBEENTANSFLNRYIVEOPKLPMDGLFLIQTAGMLQALAKQ 408
      252 VLSSSRGTTVGVGVCATDRDEPDMTRKLSIILOQTPRSP--GLFSVHSTGVITVSH 309
      409 SLKQDTPQYNLTIEVSDKPK-----TLCFVQVINVDINDQIPFEKSDYGNLTALAE 461
      310 YLDREVVDKXSLIMKVQMDGQFGILGSTSCI--ITVDSNDNAPFRQNAVE--AFVE 365
      462 DTNIGSTILTQATDADEPFTGSSKILYHIKDSBGLGVDPDPTHTNTGVIYIKPLDF 521
      366 ENAFNVEILRIPIEDKDLINTANRWVNTILKGNENGHFKISTDKETNEGVLSVVKPLNY 425
      522 ETAAVSNIVFAKNEPEPLVFGV-KYNASSFAKTLTYTVNEAPOSQAHVFOAKVEDVA 580
      426 EENRQVNLIEGVNNEAPFARDIPRYALNRAVTVHVRDIDDEGEPECTPAQAVYRIKENLA 485
      581 ICFKVGNTVAKOPE---GDISYSLRGDTRGMLKIDHVTGEISVAPLUREAGSP-----Y 633

```

```

Db      486 VGSKINGYKAYDENRNGNGLRKYKLDPKGMWITIDEISGSIITSKILDEBEVEPKNELY 545
      634 RGVVAVTEVGGSSLSVSEFHLIMDVNDNPPRLAKOYTGLPFCPLSAGSLIFETDD 633
      546 NITVLADKDQDRSCTGTAAVN--IEDVNDNPPRLQGY--VVICPKPMGYTDL--AVDP 599
      694 DQHLFRGPHFTFSGSGS--LQNDWEVSKINGTHARLS-TRHNDPEERAYVVLIRINDG 750
      600 DEPVHGAP-FYFSLPNPSPEISRLMSLTKNNDTRARLSYKMGFGF--YTIPTVWD-- 654
      751 RPLEGIVSLPVTFCSCVSGSCFRPAGHQGTIPVGNV-GILITLLVIGIILAVFIR 809
      655 RAGQAMTKLIRVNLCECTHPTQCRATSRSTGVILGKMAIILAILGILFSLVLTAVCGV 714
      810 IKKDKGK 816
      715 FGATKCK 721

```

RESULT 14

```

IUDHB
desmocolin 3b precursor - human
N/Alternate names: desmosomal glycoprotein III
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 07-May-1999
C/Accession: A40390; S16464
R/Parker, A.E.; Wheeler, G.N.; Arneemann, J.; Pidsley, S.C.; Ataliotis, P.; Thomas, C.L.;
J. Biol. Chem. 266, 10438-10445, 1991
A/Title: Desmosomal glycoproteins II and III. Cadherin-like junctional molecules generated
A/Reference number: A40390; PMID:91244819; PMID:2037591
A/Accession: A40390
A/Molecule type: mRNA
A/Residues: 1-847 <PAR>
A/Cross-references: GB:X56807
A/Note: It is uncertain whether Met-1 is the initiator or whether translation is initiated
R/Buxton, R.S.; Cowin, P.; Franke, W.W.; Garrod, D.R.; Green, K.J.; King, I.A.; Koch, P.
J. Cell Biol. 121, 481-483, 1993
A/Title: Nomenclature of the desmosomal cadherins.
A/Reference number: A43032; PMID:93252984; PMID:8466729
A/Contents: annotation; nomenclature
C/Genetics:
A/Gene: GDB:DSC3; DSC2; DSC1; DS
A/Cross-references: GDB:126552; OMIM:600271
A/Map position: 18q12.1-18q12.1
C/Keywords: cadherin; cadherin repeat homology;
C/Superfamily: cadherin; cadherin repeat homology; calcium binding; cell adhesion; duplication; glycoprotein
F/1-28/Domain: signal sequence #status predicted <PRO>
F/29-135/Domain: propeptide #status predicted <PRO>
F/136-847/Product: desmocolin 3b #status predicted <EXT>
F/136-847/Domain: extracellular #status predicted <EXT>
F/138-243/Domain: cadherin repeat homology <CR1>
F/246-355/Domain: cadherin repeat homology <CR2>
F/358-471/Domain: cadherin repeat homology <CR3>
F/474-577/Domain: cadherin repeat homology <CR4>
F/578-680/Domain: cadherin repeat homology <CR5>
F/696-718/Domain: transmembrane #status predicted <TM>
F/719-847/Domain: intracellular #status predicted <INT>
F/166,392,546,629/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match      15.4%; Score 667.5; DB 1; Length 847;
Best local similarity 30.9%; Pred. No. 6,2e-34;
Matches 186; Conservative 104; Mismatches 253; Indels 59; Gaps 21;

QY      239 WKAPKVEVENVSTDPHPKITTOVRMNDPAQYSL-----VDKEKLPFPSPIDOE- 289
      137 W-APICSMENSLGPEPLFLQOYQ--SDTAQNTIYISINGPVDG--FRNLFYVERDT 192
      290 GDIYVQPLDREKDAVFAVAKDEYKSLPLSHVAVKQINNDPFCPSPVYFEV 349
      193 GNIYCTRPVREQYESFETIAFATPTDGPYPELPLIITIEDNDYVAFETETTYFTI 252
      350 QENRILGNSIGTLTAHDREENTANSFLNRYIVEQTPKLPMDGLFLIOTVAGMLQIAKOS 409

```

```

Db      253 FENCRVGTIVGQVCARDKDEPDMTRLKYSIIQV--PSPTLFQMHPPTGVITTTSSQ 310
      410 LKKQDPQVNTLTVESDKD-----FKTLQFQVNVNDINDQPIFEKSYGNLTAEEDN 464
      311 LDELIDKYOQLKIKVDMQGOYFGLOTSTCIINDVDNDHLPFTRTSY--VTSVEENT 368
      465 IGSTIITLQATDADEPFGSSKILYHIIKQDESGRLGVDPDPTHTNGVYLIKPLDPE-- 522
      369 VDVELIRVVEDDLVNTANRANNTYILKGNENGNKITYTDKATNGSVLCVKKPLAYEEK 428
      523 -----TAAVSNIVFAE--NPEELVGVKYNASSFAKTLIVTDVNEAPQSFQVFOAKY 575
      429 QQWILQIGVNNRPAFRSASPR-----SAMSTATVTVNVEDQEGECBNPIQTVRM 480
      576 SEDVALGTQVANTADPE---GLDYSYLRGTRWKLIDHYTGIFSAPLDRA--- 629
      481 KENAEVGTTSNGYKADPPTRSSSGIRYKKLDPDTMWITIDENTGSIKVFRLDREARETI 540
      630 -GSPYRVAVATEVGGSSLSVSEFHLIMDVNDNPPRLAKOYTGLPFCPLSAPSLP 688
      541 KNGIYITVLASDQGRCTGT--LGILLQDVNDNSPFLPK--KVIYICKPMTSSAIV- 595
      689 EATDDQHLFRGPHFTFSGSGS--LQNDWEVSKINGTHARLSRHTDPEERAYVVLIRI 746
      596 -AVDPDEPI-HGPPFDFLSSTSEYQRMWRRLKAINDTARLSYQN--DPEFGSYVPITV 652
      747 NDGRRPLEGIVSLPVTFCSCV-EGSCFRPAGHQGTIPVGNV-GILITLLVIGIILAV 805
      653 RD--RLGMSVSVSLDVTLCDCITENDCTRVDPRIGGGVQGLKMAIILAILGIALLPCT 710
      806 VF 807
      711 LP 712

```

RESULT 15

```

IUDHA
desmocolin 3a precursor - human
N/Alternate names: desmosomal glycoprotein II
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 07-May-1999
C/Accession: B40390; S16465
R/Parker, A.E.; Wheeler, G.N.; Arneemann, J.; Pidsley, S.C.; Ataliotis, P.; Thomas, C.L.;
J. Biol. Chem. 266, 10438-10445, 1991
A/Title: Desmosomal glycoproteins II and III. Cadherin-like junctional molecules genera
A/Reference number: A40390; PMID:91244819; PMID:2037591
A/Accession: B40390
A/Molecule type: mRNA
A/Residues: 1-901 <PAR>
A/Cross-references: GB:X56807
A/Note: It is uncertain whether Met-1 is the initiator or whether translation is initiated
R/Buxton, R.S.; Cowin, P.; Franke, W.W.; Garrod, D.R.; Green, K.J.; King, I.A.; Koch, P.
J. Cell Biol. 121, 481-483, 1993
A/Title: Nomenclature of the desmosomal cadherins.
A/Reference number: A43032; PMID:93252984; PMID:8466729
A/Contents: annotation; nomenclature
C/Genetics:
A/Gene: GDB:DSC3; DSC2; DSC1; DS
A/Cross-references: GDB:126552; OMIM:600271
A/Map position: 18q12.1-18q12.1
C/Keywords: cadherin; cadherin repeat homology;
C/Superfamily: cadherin; cadherin repeat homology; calcium binding; cell adhesion; duplication; glycoprotein
F/1-28/Domain: signal sequence #status predicted <PRO>
F/29-135/Domain: propeptide #status predicted <PRO>
F/136-901/Product: desmocolin 3a #status predicted <EXT>
F/136-901/Domain: extracellular #status predicted <EXT>
F/138-243/Domain: cadherin repeat homology <CR1>
F/246-355/Domain: cadherin repeat homology <CR2>
F/358-471/Domain: cadherin repeat homology <CR3>
F/474-577/Domain: cadherin repeat homology <CR4>
F/578-680/Domain: cadherin repeat homology <CR5>
F/696-718/Domain: transmembrane #status predicted <TM>

```

F:719-901/Domain: intracellular #status predicted <INT>
F:166 392,546,629/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:664/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match

15.4%; Score 667.5; DB 1; Length 901;

Best Local Similarity 30.9%; Pred. No. 6,9e-34;
Matches 186; Conservative 104; Mismatches 253; Indels 59; Gaps 21;

```
QY 239 WKAKPVMVWNSNDPPHRIKITYRWNDDPGAQYSL-----VDKEKLPFPFSIDOE- 289
DB 137 W-APICSMENSLGPPPLFQQVQ-SDTAQNITYISIRGPVQDE--PRNLFYVERDT 192
QY 290 GDIYVTOPLDREKDAVYFVAADKDEKPLSLYPLHVKYKDINDNPTCPSPVTFEV 349
DB 193 GNLACTRPVDRQYESFELLIAFTTPDGYTPELPLIKIEDENDNVPITFEETFTI 252
QY 350 QENRNLNSIGTLTAHNDRENTANSFLNRIVEQTPKLPMDGLFLIQTVAQMLQAKOS 409
DB 253 FENCRVGTWQVQCATKDEBDTWHTRLKYSIIQVQ--PSPTLFSMHPTGVITTSQ 310
QY 410 LKKQDTQVNLITIEVSKD-----FKTLCFQVQINVIDINDQIPIFEKSDVGNLTLADTN 464
DB 311 LDRELIRKYQKIKIVQMDQYFGLQTTSTCIINDVDNHLPTFTKTSY--VTSVEBNT 368
QY 465 IGSTILITQATDADEPFTGSSKILYHIKDSERLGVTDPTNTGYVLIKPLDPE-- 522
DB 369 VDVETILKRVTEKDLVNTAMRANYTIIKNGENGFKIVDAKNEGVLGVKPLANEK 428
QY 523 -----TAVSNIVFKAE-NPEPLVFGVKYNASSPAKFTLIVTDVNEAPQFSQHVFOAKV 575
DB 429 QQMILQIGVNVAEAFESREASPR-----SAMSTATVTVNVEDQDEGPECNPIQTVRM 480
QY 576 SEDVAIGTKVGNVAKDPE---GLDISLNGDTRGMKIDHVTGEIIFSVAFLDREA--- 629
DB 481 KENAEVGTSTNGVAYADETSSSGIRYKLTDPGTWTIDENTGSIKVFSLDREAETI 540
QY 630 -GSPYRQVAVATEVGGSSLSVSEFHLIMDNDNPPRLAKDYTGLFECPLSAPGSLIF 688
DB 541 KNGIYNITVLASDQGRCTGT--LGITLDVNDNSPPIPK--KTVIICKPTMSAEIV- 595
QY 689 EATDDQHLFRGPHFTFSLSGS--LQNDWEVSKINGTHARLSTRHTDFFERAYVVLIRI 746
DB 596 -AVDPDEPI-HGPPFDFFSLSTSEVQRMWRKAIINDTAARLSYQN-DPPFGSYVVPITV 652
QY 747 NDGRRPPLLEGIVSLPYTFCSGV-EGSCFRPAHQGTGPTVGMAYGILLTLLVGIILAV 805
DB 653 RD--RLGMSVYSLDVTLCDCITENDCTHRVDPRIGGGVQLGKMAITAILGLALFCI 710
QY 806 VF 807
DB 711 LF 712
```

Search completed: June 22, 2003, 01:08:45
Job time : 47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 01:04:45 ; Search time 62 Seconds
(without alignments)
1452.066 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIIQAHLSICLIMLYATG.....DKGKNVSAQAEVPLRS 832

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCRUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUB pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4321	100.0	832	9	US-10-025-380-1081
2	4321	100.0	832	10	US-09-922-217-1081
3	4321	100.0	832	10	US-09-833-263-1081
4	4208	97.4	958	9	US-10-025-380-1087
5	4208	97.4	958	10	US-09-922-217-1087
6	4208	97.4	958	10	US-09-833-263-1087
7	908.5	21.0	807	9	US-10-063-547-98
8	908.5	21.0	807	9	US-10-063-516-98
9	908.5	21.0	807	9	US-10-063-502-98
10	908.5	21.0	807	9	US-10-227-884-188
11	908.5	21.0	807	9	US-10-230-163-188
12	908.5	21.0	807	9	US-10-006-856A-229
13	908.5	21.0	807	9	US-10-218-631-188
14	908.5	21.0	807	9	US-10-230-338-188
15	908.5	21.0	807	9	US-10-063-518-98
16	908.5	21.0	807	9	US-10-063-598-98
17	908.5	21.0	807	9	US-10-227-693-98
18	908.5	21.0	807	9	US-10-230-414-188
19	908.5	21.0	807	9	US-10-006-818A-229

20	908.5	21.0	807	9	US-10-015-393A-229	Sequence 229, App
21	908.5	21.0	807	9	US-10-063-567-98	Sequence 98, App1
22	908.5	21.0	807	9	US-10-216-159A-188	Sequence 188, App
23	908.5	21.0	807	9	US-09-846-374-229	Sequence 229, App
24	908.5	21.0	807	9	US-10-012-121A-229	Sequence 229, App
25	908.5	21.0	807	9	US-10-015-869A-229	Sequence 229, App
26	908.5	21.0	807	9	US-10-063-538-98	Sequence 98, App1
27	908.5	21.0	807	9	US-10-218-849-188	Sequence 188, App
28	908.5	21.0	807	9	US-10-227-873-188	Sequence 188, App
29	908.5	21.0	807	9	US-10-227-883-188	Sequence 188, App
30	908.5	21.0	807	9	US-10-063-599-98	Sequence 98, App1
31	908.5	21.0	807	9	US-10-219-076-188	Sequence 188, App
32	908.5	21.0	807	9	US-10-230-434-188	Sequence 188, App
33	908.5	21.0	807	9	US-10-006-116A-229	Sequence 229, App
34	908.5	21.0	807	9	US-10-006-117A-229	Sequence 229, App
35	908.5	21.0	807	9	US-10-013-913A-229	Sequence 229, App
36	908.5	21.0	807	9	US-10-017-527A-229	Sequence 229, App
37	908.5	21.0	807	9	US-10-063-595-98	Sequence 98, App1
38	908.5	21.0	807	9	US-10-219-003-188	Sequence 188, App
39	908.5	21.0	807	9	US-10-219-075-188	Sequence 188, App
40	908.5	21.0	807	9	US-10-219-464-188	Sequence 188, App
41	908.5	21.0	807	9	US-10-219-466-188	Sequence 188, App
42	908.5	21.0	807	9	US-10-219-479-188	Sequence 188, App
43	908.5	21.0	807	9	US-10-219-481-188	Sequence 188, App
44	908.5	21.0	807	9	US-10-230-260-188	Sequence 188, App
45	908.5	21.0	807	9	US-10-232-231-188	Sequence 188, App

ALIGNMENTS

RESULT 1
US-10-025-380-1081
Sequence 1081, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeline Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tongtong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121, 471C14
CURRENT APPLICATION NUMBER: US/10/025,380
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1081
LENGTH: 832
TYPE: PRT
ORGANISM: Homo sapiens
US-10-025-380-1081

Query Match 100.0%; Score 4321; DB %; Length 832;

Best Local Similarity 100.0%; Pred. No. 1.6e-315; Mismatches 0; Indels 0; Gaps 0;

DB 1 MIIQAHLSICLIMLYATGYGEGKFGSLKPMFTSIIYEGEPSSQIIIFQXANPAVTF 60
1 MIIQAHLSICLIMLYATGYGEGKFGSLKPMFTSIIYEGEPSSQIIIFQXANPAVTF 60

```

QY 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVKDIN 120
DB 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVKDIN 120
QY 121 DNRPTFLOSKEYESVRONSRRPGKPEFLYYNATLDDPATNGQLYYQIVIQLPMINNMYF 180
DB 121 DNRPTFLOSKEYESVRONSRRPGKPEFLYYNATLDDPATNGQLYYQIVIQLPMINNMYF 180
QY 181 QINNKGAISLTREGSOELNPAKPSYNLVISYKDMGGSNSFSDTTSVDIIVTENIMK 240
DB 181 QINNKGAISLTREGSOELNPAKPSYNLVISYKDMGGSNSFSDTTSVDIIVTENIMK 240
QY 241 APRKVENWENSTDPHPKITQVRMNDGAQYSLVDKEXLPRPFPSIDOEGLIYVTOPLDR 300
DB 241 APRKVENWENSTDPHPKITQVRMNDGAQYSLVDKEXLPRPFPSIDOEGLIYVTOPLDR 300
QY 301 EKDQAVFYAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTFVEVOENERLGNISIG 360
DB 301 EKDQAVFYAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTFVEVOENERLGNISIG 360
QY 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLIQTAYAGMLQAKOSLKKODTPOYNL 420
DB 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLIQTAYAGMLQAKOSLKKODTPOYNL 420
QY 421 TIEVSDKDFKLCFVQINVDINDQIPFEKSDYGNLTAEADNIGSTILTIOATDADER 480
DB 421 TIEVSDKDFKLCFVQINVDINDQIPFEKSDYGNLTAEADNIGSTILTIOATDADER 480
QY 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFETAASNIYFKAENPEPLV 540
DB 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFETAASNIYFKAENPEPLV 540
QY 541 FGKYNASSPAKFTLLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGGLDISY 600
DB 541 FGKYNASSPAKFTLLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGGLDISY 600
QY 601 SLRGDTRGMLKIDHTGELFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLIMDV 660
DB 601 SLRGDTRGMLKIDHTGELFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLIMDV 660
QY 661 NDNPRLAKDYTGLFCHPLSAPGSLIFEATDDQHLFRGPHFTFSLGSSLQNDWEVSK 720
DB 661 NDNPRLAKDYTGLFCHPLSAPGSLIFEATDDQHLFRGPHFTFSLGSSLQNDWEVSK 720
QY 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLLEGIVSLPVTFCSCEGSCFRPAGHOT 780
DB 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLLEGIVSLPVTFCSCEGSCFRPAGHOT 780
QY 781 GIPVGNAGVILITLLVIGIILAVVFIRIKKDKGKNVESAOASEVKPLRS 832
DB 781 GIPVGNAGVILITLLVIGIILAVVFIRIKKDKGKNVESAOASEVKPLRS 832

```

RESULT 2

```

US-09-922-217-1081
; Sequence 1081, Application US/09922217
; Patent No. US2002007641A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secretat, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13

```

```

; CURRENT APPLICATION NUMBER: US/09/922,217
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1081
; LENGTH: 832
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-922-217-1081

```

```

Query Match 100.0%; Score 4321; DB 10; Length 832;
Best Local Similarity 100.0%; Pred. No. 1,6e-315;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MIIQAHLSICLMLYLANGYGOEGKSPGLKMTESIYEGSPSIIIFQKPNPAVFE 60
DB 1 MIIQAHLSICLMLYLANGYGOEGKSPGLKMTESIYEGSPSIIIFQKPNPAVFE 60
QY 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVKDIN 120
DB 61 ELTGETDNIFVIEREGLLYYNRALDRETRSTHNLQVAALDANGIIVEGPVITIEVKDIN 120
QY 121 DNRPTFLOSKEYESVRONSRRPGKPEFLYYNATLDDPATNGQLYYQIVIQLPMINNMYF 180
DB 121 DNRPTFLOSKEYESVRONSRRPGKPEFLYYNATLDDPATNGQLYYQIVIQLPMINNMYF 180
QY 181 QINNKGAISLTREGSOELNPAKPSYNLVISYKDMGGSNSFSDTTSVDIIVTENIMK 240
DB 181 QINNKGAISLTREGSOELNPAKPSYNLVISYKDMGGSNSFSDTTSVDIIVTENIMK 240
QY 241 APRKVENWENSTDPHPKITQVRMNDGAQYSLVDKEXLPRPFPSIDOEGLIYVTOPLDR 300
DB 241 APRKVENWENSTDPHPKITQVRMNDGAQYSLVDKEXLPRPFPSIDOEGLIYVTOPLDR 300
QY 301 EKDQAVFYAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTFVEVOENERLGNISIG 360
DB 301 EKDQAVFYAVAKDEYKPLSYPLEIHVKVDINDNPPTCPSPVTFVEVOENERLGNISIG 360
QY 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLIQTAYAGMLQAKOSLKKODTPOYNL 420
DB 361 TLTAHDEDEENTANSFLNRYIVEQTPKLPMDGFLIQTAYAGMLQAKOSLKKODTPOYNL 420
QY 421 TIEVSDKDFKLCFVQINVDINDQIPFEKSDYGNLTAEADNIGSTILTIOATDADER 480
DB 421 TIEVSDKDFKLCFVQINVDINDQIPFEKSDYGNLTAEADNIGSTILTIOATDADER 480
QY 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFETAASNIYFKAENPEPLV 540
DB 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIKKPLDEFETAASNIYFKAENPEPLV 540
QY 541 FGKYNASSPAKFTLLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGGLDISY 600
DB 541 FGKYNASSPAKFTLLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGGLDISY 600
QY 601 SLRGDTRGMLKIDHTGELFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLIMDV 660
DB 601 SLRGDTRGMLKIDHTGELFSVAPLDRAGSPRYOVVATEVGGSSLSVSEFHLIMDV 660
QY 661 NDNPRLAKDYTGLFCHPLSAPGSLIFEATDDQHLFRGPHFTFSLGSSLQNDWEVSK 720
DB 661 NDNPRLAKDYTGLFCHPLSAPGSLIFEATDDQHLFRGPHFTFSLGSSLQNDWEVSK 720
QY 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLLEGIVSLPVTFCSCEGSCFRPAGHOT 780
DB 721 INGTARLSTRHTDPEERAYVVLIRINDGRPLLEGIVSLPVTFCSCEGSCFRPAGHOT 780
QY 781 GIPVGNAGVILITLLVIGIILAVVFIRIKKDKGKNVESAOASEVKPLRS 832
DB 781 GIPVGNAGVILITLLVIGIILAVVFIRIKKDKGKNVESAOASEVKPLRS 832

```

RESULT 3

```

US-09-933-263-1081

```

```

; Sequence 1081, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stoik, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1081
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1081

```

```

Query Match      100.0%; Score 4321; DB 10; Length 832;
Best Local Similarity 100.0%; Pred. No. 1,6e-315;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MIIQAHLSLCLMLYATGYGSGKSGPLKMTFSIYEGQEPSQIIFQKPNPAVTF 60
DB 1 MIIQAHLSLCLMLYATGYGSGKSGPLKMTFSIYEGQEPSQIIFQKPNPAVTF 60
QY 61 ELTGEDNITFIRREGLLYNNRDLRSTRTHNLQVVALDANGIIVGPPITIEVDIN 120
DB 61 ELTGEDNITFIRREGLLYNNRDLRSTRTHNLQVVALDANGIIVGPPITIEVDIN 120
QY 121 DNRPTFLOSKEYGSRVQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYF 180
DB 121 DNRPTFLOSKEYGSRVQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYF 180
QY 121 DNRPTFLOSKEYGSRVQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYF 180
DB 121 DNRPTFLOSKEYGSRVQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYF 180
QY 181 QINNKTGAISLTREGSELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMK 240
DB 181 QINNKTGAISLTREGSELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMK 240
QY 241 APPKVENENSTDPHPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDR 300
DB 241 APPKVENENSTDPHPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDR 300
QY 241 APPKVENENSTDPHPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDR 300
DB 241 APPKVENENSTDPHPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDR 300
QY 301 EEKDAVYFAVADEYKPLSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIG 360
DB 301 EEKDAVYFAVADEYKPLSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIG 360
QY 301 EEKDAVYFAVADEYKPLSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIG 360
DB 301 EEKDAVYFAVADEYKPLSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIG 360
QY 361 TLTAHDEBENTANSFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNL 420
DB 361 TLTAHDEBENTANSFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNL 420
QY 361 TLTAHDEBENTANSFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNL 420
DB 361 TLTAHDEBENTANSFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNL 420
QY 421 TIEVSDKDFKTLCFVQINVIDINDQIPIFEKSDYGNITLADBNIGSTIITLQATDADER 480
DB 421 TIEVSDKDFKTLCFVQINVIDINDQIPIFEKSDYGNITLADBNIGSTIITLQATDADER 480
QY 421 TIEVSDKDFKTLCFVQINVIDINDQIPIFEKSDYGNITLADBNIGSTIITLQATDADER 480
DB 421 TIEVSDKDFKTLCFVQINVIDINDQIPIFEKSDYGNITLADBNIGSTIITLQATDADER 480
QY 481 FTGSSKLIYHIIGDSSEGRIGVDDPHTNTGYVITIKKPLDEPTAAVNIYFKAKENPRLV 540
DB 481 FTGSSKLIYHIIGDSSEGRIGVDDPHTNTGYVITIKKPLDEPTAAVNIYFKAKENPRLV 540
QY 481 FTGSSKLIYHIIGDSSEGRIGVDDPHTNTGYVITIKKPLDEPTAAVNIYFKAKENPRLV 540
DB 481 FTGSSKLIYHIIGDSSEGRIGVDDPHTNTGYVITIKKPLDEPTAAVNIYFKAKENPRLV 540
QY 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGLDISY 600
DB 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGLDISY 600
QY 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGLDISY 600
DB 541 FGKYNASSFAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTKVGNVTADEGLDISY 600
QY 601 SLNGDTRGMLKIDHVTGEISVAPLDBEAGSPYRVQVATEVGGSSISVSEPHLIMDY 660
DB 601 SLNGDTRGMLKIDHVTGEISVAPLDBEAGSPYRVQVATEVGGSSISVSEPHLIMDY 660
QY 601 SLNGDTRGMLKIDHVTGEISVAPLDBEAGSPYRVQVATEVGGSSISVSEPHLIMDY 660
DB 601 SLNGDTRGMLKIDHVTGEISVAPLDBEAGSPYRVQVATEVGGSSISVSEPHLIMDY 660
QY 661 NDNPRLAKDYTGAFCHPLSAPGSLIFBATDDQHLFRGPHFTFSLGSSSLONDMEVSK 720
DB 661 NDNPRLAKDYTGAFCHPLSAPGSLIFBATDDQHLFRGPHFTFSLGSSSLONDMEVSK 720
QY 661 NDNPRLAKDYTGAFCHPLSAPGSLIFBATDDQHLFRGPHFTFSLGSSSLONDMEVSK 720
DB 661 NDNPRLAKDYTGAFCHPLSAPGSLIFBATDDQHLFRGPHFTFSLGSSSLONDMEVSK 720
QY 721 INGTARLSTRHTDFEBRAYVVLIRINDGSRPPLGIVSLPVTFCSCVEGSCRRPAGHOT 780
DB 721 INGTARLSTRHTDFEBRAYVVLIRINDGSRPPLGIVSLPVTFCSCVEGSCRRPAGHOT 780

```

```

DB 721 INGTARLSTRHTDFEBRAYVVLIRINDGSRPPLGIVSLPVTFCSCVEGSCRRPAGHOT 780
QY 781 GIFTVGAVGILITLTLVIGIILAAVFIKXKXGNDVNSAASEVKPLRS 832
DB 781 GIFTVGAVGILITLTLVIGIILAAVFIKXKXGNDVNSAASEVKPLRS 832

```

RESULT 4

US-10-025-380-1087

```

; Sequence 1087, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1087
; LENGTH: 958
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-025-380-1087

```

```

Query Match      97.4%; Score 4208; DB 9; Length 958;
Best Local Similarity 99.4%; Pred. No. 5.9e-307;
Matches 813; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

```

```

QY 17 LANG-YGQKSGSGPLKMTFSIYEGQEPSQIIFQKPNPAVTELTGETNIFYIER 74
DB 17 LANG-YGQKSGSGPLKMTFSIYEGQEPSQIIFQKPNPAVTELTGETNIFYIER 74
QY 141 LASGSAQAQEGKSGPLKMTFSIYEGQEPSQIIFQKPNPAVTELTGETNIFYIER 200
DB 141 LASGSAQAQEGKSGPLKMTFSIYEGQEPSQIIFQKPNPAVTELTGETNIFYIER 200
QY 75 EGLLYNNRDLRSTRTHNLQVVALDANGIIVGPPITIEVDINDNPTFLOSKEYS 134
DB 75 EGLLYNNRDLRSTRTHNLQVVALDANGIIVGPPITIEVDINDNPTFLOSKEYS 134
QY 201 EGLLYNNRDLRSTRTHNLQVVALDANGIIVGPPITIEVDINDNPTFLOSKEYS 260
DB 201 EGLLYNNRDLRSTRTHNLQVVALDANGIIVGPPITIEVDINDNPTFLOSKEYS 260
QY 135 VRQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYQINNKTGAISLTRE 194
DB 135 VRQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYQINNKTGAISLTRE 194
QY 261 VRQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYQINNKTGAISLTRE 320
DB 261 VRQNSRPGKPLVYNATDLDPPATPGOLYYQIVOLPMINNMVYQINNKTGAISLTRE 320
QY 195 GSGELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMKAPKVENENSTDP 254
DB 195 GSGELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMKAPKVENENSTDP 254
QY 321 GSGELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMKAPKVENENSTDP 380
DB 321 GSGELNPAKNPSYNLVISVDMGQSGENSESDTTSVDIIVTENIMKAPKVENENSTDP 380
QY 255 HPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDBEEDKDAVYFAVAKD 314
DB 255 HPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDBEEDKDAVYFAVAKD 314
QY 381 HPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDBEEDKDAVYFAVAKD 440
DB 381 HPKIKITQVRWMDGAQYSLVDKELRFPFSDIOEGDIYVTOPLDBEEDKDAVYFAVAKD 440
QY 315 EYKSPSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIGTLTAHDEBENTAN 374
DB 315 EYKSPSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIGTLTAHDEBENTAN 374
QY 441 EYKSPSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIGTLTAHDEBENTAN 500
DB 441 EYKSPSYPLEIHVKYKDIINDNPTCPSPTVFEVQENERLGNISIGTLTAHDEBENTAN 500
QY 375 SFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNLTIEVSDKDFKTLCF 434
DB 375 SFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNLTIEVSDKDFKTLCF 434
QY 501 SFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNLTIEVSDKDFKTLCF 560
DB 501 SFLNRYIVEQTPKLPMQDGLFIQTVAGMLQAKOSLKKODTPQYNLTIEVSDKDFKTLCF 560

```

QY 435 VOINVINDINQPIFEKSDYGNLTAAEDTNIGSTILTIOATDADEPFTGSSKILYHIK 494
DB 561 VOINVINDINQPIFEKSDYGNLTAAEDTNIGSTILTIOATDADEPFTGSSKILYHIK 620
QY 495 DSEGRGLVDTDPHTNTGYVLIKKPLDETAASNIIVKAEKNEPELVGVKNASSPAKFT 554
DB 621 DSEGRGLVDTDPHTNTGYVLIKKPLDETAASNIIVKAEKNEPELVGVKNASSPAKFT 680
QY 555 LIYTVNEAPQFSGHVFQAKVSESDVAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 614
DB 681 LIYTVNEAPQFSGHVFQAKVSESDVAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 740
QY 615 VTGEIFSVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRKAKDYTL 674
DB 741 VTGEIFSVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRKAKDYTL 800
QY 675 FFCGPLAPGSLIFEATDDOHLFRGPHFTFSLGSSSLQNDMEVSKINGTHARLSTHTD 734
DB 801 FFCGPLAPGSLIFEATDDOHLFRGPHFTFSLGSSSLQNDMEVSKINGTHARLSTHTD 860
QY 735 FEERAYVVLIRINDGRPPLEGIVSLPYTFCSCEVSCCFRPAHQGTGIPVGNVAVGILTT 794
DB 861 FEERAYVVLIRINDGRPPLEGIVSLPYTFCSCEVSCCFRPAHQGTGIPVGNVAVGILTT 920
QY 795 TLVIGIILAVFIRIKKDKGNVESAOASEVPLRS 832
DB 921 TLVIGIILAVFIRIKKDKGNVESAOASEVPLRS 958

RESULT 5

US-09-922-217-1087
Sequence 1087, Application US/09922217
Patent No. US20020076414A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secret, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeline Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Tonglong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole Lynn
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aljun
APPLICANT: Clapper, Jonathan D.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
FILE REFERENCE: 210121.471C13
CURRENT APPLICATION NUMBER: US/09/922.217
NUMBER OF SEQ ID NOS: 1124
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1087
LENGTH: 958
TYPE: PRT
ORGANISM: Homo sapiens
US-09-922-217-1087

Query Match 97.4%; Score 4208; DB 10; Length 958;
Best Local Similarity 99.4%; Pred. No. 5.9e-307;
Matches 813; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 17 LATG--YGOEGKFGPLKMTFSIYEGEPQIIFOFKANPAYTFELTGETDNI FYIER 74
DB 141 LASSGAAQBEKFGPLKMTFSIYEGEPQIIFOFKANPAYTFELTGETDNI FYIER 200
QY 75 EGLIYNBALDRETRSTNLOVALDANGIIVBGPVPTTIYKQINDNRPFLQSKYEGS 134
DB 201 EGLIYNBALDRETRSTNLOVALDANGIIVBGPVPTTIYKQINDNRPFLQSKYEGS 260
QY 135 VRQNSRPGKFLYVNAATDLDPAIPNGQLYQVIGLPMINWVFOINNTGAI SLTRE 194

DB 261 VRQNSRPGKFLYVNAATDLDPAIPNGQLYQVIGLPMINWVFOINNTGAI SLTRE 320
QY 195 GSGELNPAKPSNLVYSKDMGQSENSFSDTSDIIVTENIMKAPKPEVENS TD 254
DB 321 GSGELNPAKPSNLVYSKDMGQSENSFSDTSDIIVTENIMKAPKPEVENS TD 380
QY 255 HPKITQVRWNDGAQYSLVDEKELPRFPESIDQEGDIYVTOPLDREKDAVYFAVAKD 314
DB 381 HPKITQVRWNDGAQYSLVDEKELPRFPESIDQEGDIYVTOPLDREKDAVYFAVAKD 440
QY 315 EYKPLSYPLEIHVKYKINDNPTCPSPVYVEVOENERLQNSIGTLTAHDEENTAN 374
DB 441 EYKPLSYPLEIHVKYKINDNPTCPSPVYVEVOENERLQNSIGTLTAHDEENTAN 500
QY 375 SFNANYIVEQTPPLPMDGLFLIOTYAGMLQAKQSLKKQDTPQYNLTIEVSDQFTLCF 434
DB 501 SFNANYIVEQTPPLPMDGLFLIOTYAGMLQAKQSLKKQDTPQYNLTIEVSDQFTLCF 560
QY 435 VOINVINDINQPIFEKSDYGNLTAAEDTNIGSTILTIOATDADEPFTGSSKILYHIK 494
DB 561 VOINVINDINQPIFEKSDYGNLTAAEDTNIGSTILTIOATDADEPFTGSSKILYHIK 620
QY 495 DSEGRGLVDTDPHTNTGYVLIKKPLDETAASNIIVKAEKNEPELVGVKNASSPAKFT 554
DB 621 DSEGRGLVDTDPHTNTGYVLIKKPLDETAASNIIVKAEKNEPELVGVKNASSPAKFT 680
QY 555 LIYTVNEAPQFSGHVFQAKVSESDVAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 614
DB 681 LIYTVNEAPQFSGHVFQAKVSESDVAIGTKVGNVTAADPEGLDISYSLRGDTGWLKIDH 740
QY 615 VTGEIFSVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRKAKDYTL 674
DB 741 VTGEIFSVAPLDRAGSPYRVQVAVTEVGSLSVSEFHLIMDVNDNPRKAKDYTL 800
QY 735 FEERAYVVLIRINDGRPPLEGIVSLPYTFCSCEVSCCFRPAHQGTGIPVGNVAVGILTT 794
DB 861 FEERAYVVLIRINDGRPPLEGIVSLPYTFCSCEVSCCFRPAHQGTGIPVGNVAVGILTT 920
QY 795 TLVIGIILAVFIRIKKDKGNVESAOASEVPLRS 832
DB 921 TLVIGIILAVFIRIKKDKGNVESAOASEVPLRS 958

RESULT 6

US-09-833-263-1087
Sequence 1087, Application US/09833263
Patent No. US20020110547A1
GENERAL INFORMATION:
APPLICANT: Wang, Aljun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Stolk, John A.
APPLICANT: Meagher, Madeline J.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121.471C12
CURRENT APPLICATION NUMBER: US/09/833.263
NUMBER OF SEQ ID NOS: 1093
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1087
LENGTH: 958
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-263-1087

Query Match 97.4%; Score 4208; DB 10; Length 958;
Best Local Similarity 99.4%; Pred. No. 5.9e-307;
Matches 813; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

```

QY 17 LATG--YGOEGKSGPLKPMTPFSIYEGQSPQIIFQFANPAPVTFELTGTNDIPIYER 74
D 141 LASGSAAOEGKSGPLKPMTPFSIYEGQSPQIIFQFANPAPVTFELTGTNDIPIYER 200
QY 75 EGLLYNRALDRERSTHNLQVALDANGIIVEGVPVTTIEVKDINDRPFLOSKEYS 134
D 201 EGLLYNRALDRERSTHNLQVALDANGIIVEGVPVTTIEVKDINDRPFLOSKEYS 260
QY 135 VRQNSRPGKPELYVNAVDLDDPATPNQLYQYQIYQIOLPMINWYFOINNTGTAISLTRE 194
D 261 VRQNSRPGKPELYVNAVDLDDPATPNQLYQYQIYQIOLPMINWYFOINNTGTAISLTRE 320
QY 195 GSQGLNPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTENIWKAPKPEWENSTDP 254
D 321 GSQGLNPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTENIWKAPKPEWENSTDP 380
QY 255 HPKITQVRNMDPQAQYSLVDKEKLPFRPFSIDQEGDIYVTQPLDREKDAYVYVAVAKD 314
D 381 HPKITQVRNMDPQAQYSLVDKEKLPFRPFSIDQEGDIYVTQPLDREKDAYVYVAVAKD 440
QY 315 EYGPRLSYPLIEHYKVDINDNPTCPSPVTVFEQENRERLNSIGTLTAHDEENNTAN 374
D 441 EYGPRLSYPLIEHYKVDINDNPTCPSPVTVFEQENRERLNSIGTLTAHDEENNTAN 500
QY 375 SFLNRYLVEQTPKLPMDGLFIQTYAGMLQAKQSLKKODTPQYNTLIEVSDKDFKTLCE 434
D 501 SFLNRYLVEQTPKLPMDGLFIQTYAGMLQAKQSLKKODTPQYNTLIEVSDKDFKTLCE 560
QY 435 VOINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLIQAOTDADDEPTGSSKILYHILKG 494
D 561 VOINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLIQAOTDADDEPTGSSKILYHILKG 620
QY 495 DSEGRGLVDTPDHTNTGYVIIKKPLDEPETAANSIVFEKAENPEPLVGVKNASSPAKFT 554
D 621 DSEGRGLVDTPDHTNTGYVIIKKPLDEPETAANSIVFEKAENPEPLVGVKNASSPAKFT 680
QY 555 LIYTDVNEAPQFSQHFQAKVSEDAVIGTVGNVTAKDPEGLDISYSLRGDTRGWLKIDH 614
D 681 LIYTDVNEAPQFSQHFQAKVSEDAVIGTVGNVTAKDPEGLDISYSLRGDTRGWLKIDH 740
QY 615 VTGEITSVAPLDRBAGSPYVQVAVTEVGGSSLSVSSEFHLIMDVNDNPPRLAKDTGL 674
D 741 VTGEITSVAPLDRBAGSPYVQVAVTEVGGSSLSVSSEFHLIMDVNDNPPRLAKDTGL 800
QY 675 FFCGHPASPGSLIFEATDDOHLFRGPHFTFSLGSGSLQNDMEVSKINGTHARLSTRHTD 734
D 801 FFCGHPASPGSLIFEATDDOHLFRGPHFTFSLGSGSLQNDMEVSKINGTHARLSTRHTD 860
QY 735 FEBRAYVYLIRINDGRRPLEGIVSLPVTECSCEVSGCFRRPAGHOTGIPVGNVAVGILLT 794
D 861 FEBRAYVYLIRINDGRRPLEGIVSLPVTECSCEVSGCFRRPAGHOTGIPVGNVAVGILLT 920
QY 795 TLLVIGIILAVVIRIKKQKGNVSAQASEVKAFLRS 832
D 921 TLLVIGIILAVVIRIKKQKGNVSAQASEVKAFLRS 958

```

```

RESULT 7
US-10-063-547-98
; Sequence 98, Application US/10063547
; Publication No. US20020182638A1
; GENERAL INFORMATION:
; APPLICANT: Baton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerltsen, Mary B.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Guiney, Austin L.
; APPLICANT: Matanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,547
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 98
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-547-98

```

```

Query Match 21.0%; Score 908.5; DB 9; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

```

```

QY 62 LTGE---TDNITVIERE--GLLYNRALDRERSTHNLQVALDANGIIVEGVPVTTIEV 116
D 57 LSGDSGKATEGPPAMPDPSGFLVTRALDREQAQYQVLTLEMQGHVLMGPOVLYVHV 116
QY 117 KDINDRPFLOSKEYSVRQNSRPGKPELYVNAVDLDDPATPNQLYQYQIOLPMINN 176
D 117 KDENDQVPHFSQAIYARLSRGTRPGIPLFLFASDRDEGTANSDLRFHILSQAAPQS 176
QY 177 VMYFOINNTGTAISLTREGSOELNPAKNSYNLVISVKMGQSGSENSFSDTTSVDIIVTE 236
D 177 PMFOLEPRIGALALPKGSTSLDHALBERTYQLLVQVKMGDQA--SGHQATATVEVSIIE 235
QY 237 NIWKAPEVWENSTDPHPPIKITQVRNMDPQAQYSLVDKEKLPFRPFSIDQEGDIYVTQ 296
D 236 STWVSLPPIHLEMLVLYVPHMAQVHWSGQDVHYH---ESHPPPEFVNAEGNLYVTR 292
QY 297 PLDREKDAYVYVAVAKDEYKPLSYPLIEHYKVDINDNPTCPSPVTVFEQENRERLQ 356
D 293 ELDRERQAYTLQVRQNSHGEDYAPPLFLHVLVMDENNVPICPRDPVTISIPELSPRG 352
QY 357 NSIGTLTAHDEENNTANFVNYRIVEQTPKLPMDG-LFLIQTAGMLQAKQSLKK-OD 414
D 353 TEVTRLSADADAPGSPNHHVYQLSPPEQSVBERAQQVDPSSSVTLGVPLRAQGN 412
QY 415 TPQYNTLIEVSDKD--FKTLCFYQVINVIDINDQIPFEKSDYGNLTLAEDTNGSTILLI 472
D 413 ILLVLAAMDLAGAEGGFSSTCEVAVAVTDINDHAPFITSQIGISLPDVEGTVLVAML 472
QY 473 QATDAD--EPTGSSKIL-YHIKDSERGLGVDTPDHTNTGYVIIR--KPDEPETAASN 528
D 473 TAIDADLEP--AFRLMDFAIERGDTGTFGLDWEF--DSGHVRLTLCNLSYEAPASHE 527
QY 529 IVFEKAENPEPLVGVKNASSPAKFTLIYTDVNEAPQFSQHFQAKVSEDAVIGTVGNV 588
D 528 VVVVVGSAKLV--GPGPGGATATVTVLYVERVMPPKLDQESYEASVPIASAPGSFLLTI 586
QY 589 TAKDEPGLDISYSLRGDTRGWLKIDHVTGEIFSVAPLD--REAGSPYVQVAVTEVGGSSL 647
D 587 QPSDPTSRILRSLVNDSEGMLCIEKFSGEVHTAQSLOQAQGDFTVTVLVEAQDTA---- 642
QY 648 SSVSEFHLIMDVNDNPPRLADYTGILFRCHPLSAPGSLIF--EATDDOHLFRGPHFTS 706
D 643 -----LTLAVPQO-----YICTRQDHGLVSGPSKDPDASAGHP--YSPT 683
QY 707 LG-SGSLQNDMEVSKINGTHARLSTRHTDFEERAVVYLIRINDGRRPLEGIVSLPVTEC 765
D 684 LGPNPTVQGDWMLQTLNLSHAYTLTALHVEERHIIIPVYSHNQ---MMQLLRAVLYC 740
QY 766 SC-VBGSCEPRPAGHOTGIPVGNVAVGILLTLLVIGIILAVVIRIKKQKGNVSAQA 824
D 741 RCNVEGQCKRWKGMKGMPTKLSAVGIIIVGLTVAIGIFILIFTHTMTRSKKDPQAPDS 800
QY 825 SEVK 828
D 801 VPLK 804

```

RESULT 8
 US-10-063-616-98
 / Sequence 98, Application US/10063616
 / Publication No. US20030013855A1
 / GENERAL INFORMATION:
 / APPLICANT: Filvaroff, Ellen
 / APPLICANT: Eaton, Dan L.
 / APPLICANT: Gerritsen, Mary E.
 / APPLICANT: Goddard, Audrey
 / APPLICANT: Godowski, Paul J.
 / APPLICANT: Grimaldi, Christopher J.
 / APPLICANT: Gurney, Austin L.
 / APPLICANT: Matanabe, Colin K.
 / APPLICANT: Wood, William I.
 / TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 / FILE REFERENCE: P3230R1C1
 / CURRENT APPLICATION NUMBER: US/10/063,616
 / CURRENT FILING DATE: 2002-05-03
 / Prior Application removed - See File Wrapper or Palm
 / NUMBER OF SEQ ID NOS: 170
 / SEQ ID NO 98
 / LENGTH: 807
 / TYPE: PRT
 / ORGANISM: Homo Sapien
 US-10-063-616-98

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

62 LTGE---TDNIFIERE-GLYYNRALDRETRSTHNLQVLAALDANGIYEGVPITIEV 116
 57 LSGSGKATGPRPMADPSGFLVTRALDREBOAEYQVTLMOQGHVLMGPVLAHV 116
 117 KDINDNRPFLQSKYEGSVQRNSRPGKFLVYNAATLDDPATNGOLYYQIVLQPMINN 176
 117 KDENDQVPHSQAITYRARLSRGTGPGIFLELASDDEPTANSIDRPHILSQAPQPS 176
 177 VMTFOINKTGALSLTRGSOELNPAKNPSYNIYISKDGGSGENSESDTTSVDIYTE 236
 177 PDMFQLEPRGALALSPKSGTSLDHALERTYQLVQVDMGDA-SGHQATATVEVSII 235
 237 NIKAPRVEMVENSSTDPIKITYRANMDGAGYSLVDEKELRPFPSIDQEGDIYVQ 296
 236 STWVSLPFIHLAENLKVLPFHMAQVMSGADVHHL---ESHPPGEFVNAEGNLVYTR 292
 297 PLDEREKDAVYFAVAKDEYKPLSYPLEIHVYKVDINDNPTCPSPVTFVEVOENRLG 356
 293 ELDERAQAEYLLQVRAQNSHGEDYAAPLELHVLMENDNVPICPRPDPTVSIPELSPPG 352
 357 NSIGTLTAHDEENTANSFLNRIVEQTPKLPMDG-LFLIQYAGMLQAKOSLKK-QD 414
 353 TEVTRLSAEDADAGSPNSHVYQLLSPEPEDGVEGAFQVDPSTSGSVTLGLPLRAGON 412
 415 TPQYNLTIEVSKD--FKTLCFYQVINVDINDQIPFEKSDYGNLTLAEDTNISITLTI 472
 413 ILLVILAMDLAGAGGSSSTCEVEAVATDINDHAPEFTSQIGPISLPEDEVBGTLVAML 472
 473 QATDAD-EPFTGSSKIL-YHIKGSDEGRLEVDTPHTNTGYLIK--KPLDFETAASN 528
 473 TALDADLEP---ARLMDFAIERDGTBTGFLDWER--DSGHVRLRCKNLSYAAASH 527
 529 IVFAENPEPLVFGVKTNASSFAKFTLLVTVDNEAPQSGHVFQAKVSEDAVIGTKVGNV 588
 528 VVVVVVAVAKLV-GPGPGPAGATATVTVVERVMPPKLDQESYASVYISAPAGSFLITI 586
 589 TAKPREGIDISYLRGDTRGWLKIDHTGELFSAVPLD-REAGSPYRQVAVATEVGSSTL 647
 587 QPSDPISRTLAFSLVNDSEGLCIKESGSEVHTAQSLOGAQPGLTYTVLVAQDTA--- 642
 648 SSVSEFHLIIMDVNDNPRFLAKDYTGLEFCFAPLSAPGSLIF-EATDDQHLFRGPHFFS 706

Db 643 -----LTLAPVPSQ-----YLCTPRQDHGLIVSGPSKDPDLASGHP-YSFT 663
 Qy 707 LG-SGSLQNDWESKINGTHARLSTRHTDPEERAAVVLIRINDGRRPLEGIYSLPTFC 765
 Db 684 LQNPFTVQRDMRQLQTLNGSAVLTLLAHWVEPREHIIIPVVSNAQ---MQQLVRYVC 740
 Qy 766 SC-VEGSCFRPAGHQGTPIPVNAVGIILTTLLVIGIILAVERIRIKKQKDNVESAQA 824
 Db 741 RNVGEGQCRKVKQRMKMGWPTKLSAVGILVGTVAIGIFILLIPTHTWMSKKDPDPADS 800
 Qy 825 SEVX 828
 Db 801 VPLK 804

RESULT 9
 US-10-063-502-98
 / Sequence 98, Application US/10063502
 / Publication No. US20030023042A1
 / GENERAL INFORMATION:
 / APPLICANT: Eaton, Dan L.
 / APPLICANT: Filvaroff, Ellen
 / APPLICANT: Gerritsen, Mary E.
 / APPLICANT: Goddard, Audrey
 / APPLICANT: Godowski, Paul J.
 / APPLICANT: Grimaldi, Christopher J.
 / APPLICANT: Gurney, Austin L.
 / APPLICANT: Matanabe, Colin K.
 / APPLICANT: Wood, William I.
 / TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 / FILE REFERENCE: P3230R1C1
 / CURRENT APPLICATION NUMBER: US/10/063,502
 / CURRENT FILING DATE: 2002-05-01
 / Prior Application removed - See File Wrapper or Palm
 / NUMBER OF SEQ ID NOS: 170
 / SEQ ID NO 98
 / LENGTH: 807
 / TYPE: PRT
 / ORGANISM: Homo Sapien
 US-10-063-502-98

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

62 LTGE---TDNIFIERE-GLYYNRALDRETRSTHNLQVLAALDANGIYEGVPITIEV 116
 57 LSGSGKATGPRPMADPSGFLVTRALDREBOAEYQVTLMOQGHVLMGPVLAHV 116
 117 KDINDNRPFLQSKYEGSVQRNSRPGKFLVYNAATLDDPATNGOLYYQIVLQPMINN 176
 117 KDENDQVPHSQAITYRARLSRGTGPGIFLELASDDEPTANSIDRPHILSQAPQPS 176
 177 VMTFOINKTGALSLTRGSOELNPAKNPSYNIYISKDGGSGENSESDTTSVDIYTE 236
 177 PDMFQLEPRGALALSPKSGTSLDHALERTYQLVQVDMGDA-SGHQATATVEVSII 235
 237 NIKAPRVEMVENSSTDPIKITYRANMDGAGYSLVDEKELRPFPSIDQEGDIYVQ 296
 236 STWVSLPFIHLAENLKVLPFHMAQVMSGADVHHL---ESHPPGEFVNAEGNLVYTR 292
 297 PLDEREKDAVYFAVAKDEYKPLSYPLEIHVYKVDINDNPTCPSPVTFVEVOENRLG 356
 293 ELDERAQAEYLLQVRAQNSHGEDYAAPLELHVLMENDNVPICPRPDPTVSIPELSPPG 352
 357 NSIGTLTAHDEENTANSFLNRIVEQTPKLPMDG-LFLIQYAGMLQAKOSLKK-QD 414
 353 TEVTRLSAEDADAGSPNSHVYQLLSPEPEDGVEGAFQVDPSTSGSVTLGLPLRAGON 412
 415 TPQYNLTIEVSKD--FKTLCFYQVINVDINDQIPFEKSDYGNLTLAEDTNISITLTI 472
 413 ILLVILAMDLAGAGGSSSTCEVEAVATDINDHAPEFTSQIGPISLPEDEVBGTLVAML 472


```

OY 473 QATDAD-BEFTGSKTL-YHIIKGDSEGRGVDTPDHNNGVYIK--KPLDETAIVSN 528
Db 473 TAIADADLEP---AFRLMDPAIEKGDTEGTGLDMEP--DSGHVRLRLCKNLSTEAAPSH 527
OY 529 IVFKANPEPFLVPGVYKYNASSFAKFTLLVTDVNEAPQFSQHVCQAKVSEDAVIGTKVNV 588
Db 529 VVVVVQSVAKLV-GGPGPGCATATVTVLVERVMPPLDQESYEASVPISAPAGSFLTLTI 566
OY 589 TAKDPEGLDISYLRKDDTGKMLKIDHYTGHSFVAPLD-REASPFRVQVVAIVEGSSSL 644
Db 589 QPSDPLSRTRLRESLVNDSEGMWLCIEKFSGEVHQAOSIQGAQPDPTVTVLEADDTA---- 642
OY 648 SSVSEFHLILMDVNDNPPRLAKOYTGILFCHPLSAPGSLIF-BATDDQCHLFRGPHTFS 706
Db 643 -----LTLAIVPSQ-----YICTPQODHGLIVSGSKPDLASGHP-XSFT 663
OY 707 LG-SGSLQNDWEVSKINGTHARLSTHRTDPEERAYVVLIRINDGRPPLEGIVSLPYTFC 765
Db 684 LGPNFTVQGRDMLQLQTINGSHAYLTLALHWEPREHIIIPVVSHNAQ---WMQGLVRYIVC 740
OY 766 SC-VEGSCPRRPHQVGIPTVGNAGVILLTLLVIGIILAIVTRIKKQKGNVESAOA 824
Db 741 RNVSGQCKRKVKRKMGMPTKLSAVGILGTVAIGFILIPHTHWTMRSKDPDOPADS 800
OY 825 SEVK 828
Db 801 VPLK 804

RESULT 10
US-10-227-884-188
: Sequence 188, Application US/10227884
: Publication No. US20030027988A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Desnoyers, Luc
: APPLICANT: Gerritsen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godwaldt, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Auecin L.
: APPLICANT: Smith, Victoria
: APPLICANT: Stephan, Jean-Philippe F.
: APPLICANT: Watanabe, Colin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: P3530PIC79
: CURRENT APPLICATION NUMBER: US/10/227,884
: CURRENT FILING DATE: 2002-08-26
: PRIOR APPLICATION NUMBER: 10/119,480
: PRIOR FILING DATE: 2002-04-09
: PRIOR APPLICATION NUMBER: 60/059113
: PRIOR FILING DATE: 1997-09-17
: PRIOR APPLICATION NUMBER: 60/062287
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/063549
: PRIOR FILING DATE: 1997-10-28
: PRIOR APPLICATION NUMBER: 60/064103
: PRIOR FILING DATE: 1997-10-31
: PRIOR APPLICATION NUMBER: 60/069873
: PRIOR FILING DATE: 1997-12-17
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/081619
: PRIOR FILING DATE: 1998-04-15

```

1	PRIOR APPLICATION NUMBER: 60/00819555
2	PRIOR FILING DATE: 1998-04-15
3	PRIOR APPLICATION NUMBER: 60/08280404
4	PRIOR FILING DATE: 1998-04-22
5	PRIOR APPLICATION NUMBER: 60/0844441
6	PRIOR FILING DATE: 1998-05-06
7	PRIOR APPLICATION NUMBER: 60/08532232
8	PRIOR FILING DATE: 1998-05-13
9	PRIOR APPLICATION NUMBER: 60/08557979
10	PRIOR FILING DATE: 1998-05-15
11	PRIOR APPLICATION NUMBER: 60/08639292
12	PRIOR FILING DATE: 1998-05-22
13	PRIOR APPLICATION NUMBER: 60/08953232
14	PRIOR FILING DATE: 1998-06-17
15	PRIOR APPLICATION NUMBER: 60/08935388
16	PRIOR FILING DATE: 1998-06-17
17	PRIOR APPLICATION NUMBER: 60/08990505
18	PRIOR FILING DATE: 1998-06-18
19	PRIOR APPLICATION NUMBER: 60/09047272
20	PRIOR FILING DATE: 1998-06-24
21	PRIOR APPLICATION NUMBER: 60/09055757
22	PRIOR FILING DATE: 1998-06-24
23	PRIOR APPLICATION NUMBER: 60/09065931
24	PRIOR FILING DATE: 1998-06-25
25	PRIOR APPLICATION NUMBER: 60/09065959
26	PRIOR FILING DATE: 1998-06-25
27	PRIOR APPLICATION NUMBER: 60/09198222
28	PRIOR FILING DATE: 1998-07-07
29	PRIOR APPLICATION NUMBER: 60/09530202
30	PRIOR FILING DATE: 1998-08-04
31	PRIOR APPLICATION NUMBER: 60/09531188
32	PRIOR FILING DATE: 1998-08-04
33	PRIOR APPLICATION NUMBER: 60/09591666
34	PRIOR FILING DATE: 1998-08-10
35	PRIOR APPLICATION NUMBER: 60/09614646
36	PRIOR FILING DATE: 1998-08-11
37	PRIOR APPLICATION NUMBER: 60/09679212
38	PRIOR FILING DATE: 1998-08-17
39	PRIOR APPLICATION NUMBER: 60/09798666
40	PRIOR FILING DATE: 1998-08-26
41	PRIOR APPLICATION NUMBER: 60/09854444
42	PRIOR FILING DATE: 1998-08-31
43	PRIOR APPLICATION NUMBER: 60/09955666
44	PRIOR FILING DATE: 1998-09-09
45	PRIOR APPLICATION NUMBER: 60/09955988
46	PRIOR FILING DATE: 1998-09-09
47	PRIOR APPLICATION NUMBER: 60/09980303
48	PRIOR FILING DATE: 1998-09-10
49	PRIOR APPLICATION NUMBER: 60/09981111
50	PRIOR FILING DATE: 1998-09-10
51	PRIOR APPLICATION NUMBER: 60/09981212
52	PRIOR FILING DATE: 1998-09-10
53	PRIOR APPLICATION NUMBER: 60/09981616
54	PRIOR FILING DATE: 1998-09-10
55	PRIOR APPLICATION NUMBER: 60/10003888
56	PRIOR FILING DATE: 1998-09-11
57	PRIOR APPLICATION NUMBER: 60/10038555
58	PRIOR FILING DATE: 1998-09-15
59	PRIOR APPLICATION NUMBER: 60/10039030
60	PRIOR FILING DATE: 1998-09-15
61	PRIOR APPLICATION NUMBER: 60/10062777
62	PRIOR FILING DATE: 1998-09-16
63	PRIOR APPLICATION NUMBER: 60/10084848
64	PRIOR FILING DATE: 1998-09-18
65	PRIOR APPLICATION NUMBER: 60/10091919
66	PRIOR FILING DATE: 1998-09-17
67	PRIOR APPLICATION NUMBER: 60/10147777
68	PRIOR FILING DATE: 1998-09-23
69	PRIOR APPLICATION NUMBER: 60/10173738
70	PRIOR FILING DATE: 1998-09-24
71	PRIOR APPLICATION NUMBER: 60/10174111
72	PRIOR FILING DATE: 1998-09-24
73	PRIOR APPLICATION NUMBER: 60/10178666

PRIOR FILING DATE:	1998-09-25
PRIOR APPLICATION NUMBER:	60/101916
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/101922
PRIOR FILING DATE:	1998-09-24
PRIOR APPLICATION NUMBER:	60/106178
PRIOR FILING DATE:	1998-10-28
PRIOR APPLICATION NUMBER:	60/106248
PRIOR FILING DATE:	1998-10-29
PRIOR APPLICATION NUMBER:	60/106464
PRIOR FILING DATE:	1998-10-30
PRIOR APPLICATION NUMBER:	60/106905
PRIOR FILING DATE:	1998-11-03
PRIOR APPLICATION NUMBER:	60/108787
PRIOR FILING DATE:	1998-11-17
PRIOR APPLICATION NUMBER:	60/108801
PRIOR FILING DATE:	1998-11-17
PRIOR APPLICATION NUMBER:	60/108849
PRIOR FILING DATE:	1998-11-18
PRIOR APPLICATION NUMBER:	60/112422
PRIOR FILING DATE:	1998-12-15
PRIOR APPLICATION NUMBER:	60/113296
PRIOR FILING DATE:	1998-12-22
PRIOR APPLICATION NUMBER:	60/113605
PRIOR FILING DATE:	1998-12-23
PRIOR APPLICATION NUMBER:	60/113621
PRIOR FILING DATE:	1998-12-23
PRIOR APPLICATION NUMBER:	60/115558
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/115565
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/115753
PRIOR FILING DATE:	1999-01-12
PRIOR APPLICATION NUMBER:	60/119549
PRIOR FILING DATE:	1999-02-10
PRIOR APPLICATION NUMBER:	60/123618
PRIOR FILING DATE:	1999-03-10
PRIOR APPLICATION NUMBER:	60/125259
PRIOR FILING DATE:	1999-03-19
PRIOR APPLICATION NUMBER:	60/125775
PRIOR FILING DATE:	1999-03-23
PRIOR APPLICATION NUMBER:	60/126773
PRIOR FILING DATE:	1999-03-29
PRIOR APPLICATION NUMBER:	60/127887
PRIOR FILING DATE:	1999-04-05
PRIOR APPLICATION NUMBER:	60/130232
PRIOR FILING DATE:	1999-04-21
PRIOR APPLICATION NUMBER:	60/131022
PRIOR FILING DATE:	1999-04-26
PRIOR APPLICATION NUMBER:	60/131270
PRIOR FILING DATE:	1999-04-27
PRIOR APPLICATION NUMBER:	60/131293
PRIOR FILING DATE:	1999-04-27
PRIOR APPLICATION NUMBER:	60/131445
PRIOR FILING DATE:	1999-04-28
PRIOR APPLICATION NUMBER:	60/134287
PRIOR FILING DATE:	1999-05-14
PRIOR APPLICATION NUMBER:	60/140650
PRIOR FILING DATE:	1999-06-22
PRIOR APPLICATION NUMBER:	60/140723
PRIOR FILING DATE:	1999-06-22
PRIOR APPLICATION NUMBER:	60/141037
PRIOR FILING DATE:	1999-06-23
PRIOR APPLICATION NUMBER:	60/144758
PRIOR FILING DATE:	1999-07-20
PRIOR APPLICATION NUMBER:	60/145688
PRIOR FILING DATE:	1999-07-26
PRIOR APPLICATION NUMBER:	60/146222
PRIOR FILING DATE:	1999-07-28
PRIOR APPLICATION NUMBER:	60/146963
PRIOR FILING DATE:	1999-08-03
PRIOR APPLICATION NUMBER:	60/149320
PRIOR FILING DATE:	1999-08-17

[illegible]

RESULT 11
US-10-230-163-188
Sequence 188, Application US/10230163
Publication No. US2003003635A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Matanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3510PIC96
CURRENT APPLICATION NUMBER: US/10/230,163
CURRENT FILING DATE: 2002-08-28
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089552
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089558
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106505
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558

```

PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835

```

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

```

QY 62 LTGE---TDNIEVERB-GLLYNRAIDRETRSTHNLQVAALDANGIIEGVPITTEV 116
DB 57 LSGSGKATBGFAMDPDSGFLVLTALDREBQAEYQLQVLEMQDGHVWGPOPLVHV 116
QY 117 KDIUNRPTLOSTYEGSVNSRGRKPELYVNTDDDPATPGQLYVOVLPLMNN 176
DB 117 KDENDQVPHFSQALYRRLSRGTRGPFLPLEASDRPEPTANSDLRFHLSQAPOPS 176
QY 177 VMYFOINNKGAISLTREBSQELNPAKNPSYNLVYSVKMGQSGSNSPDTTSVDIYTE 236

```

```

DB 177 PDWFQLEPRIGALALSPKSGTSLDHALERTYQLLVQVKMDGQA-SGHQATATVEVSILE 235
QY 237 NIMKAPKPVEMWENSTDPHPIKITYRWMDPGAQSVLDKELPRPFSIDEGDITYTO 296
DB 236 STWVSLPETHLENKLVYPHMAQVHMSGGVHNL---ESHPPEPFEVNAEGNLYYTR 292
QY 297 PLDBEKDAVYFAVAKDEYKPLSYBLEIHKVKDINDNPPTCPSPVTFVEQENERIG 356
DB 293 ELDBENQAEYLLQVRQNSHGEDYAPLEHLVLMENDNNVPICPRPDPTVSIPELSPG 352
QY 357 NSIGTLTAHDREENTANSFLNRYIEQTPKLPMDG-LLIQTYAGMLQAKSLKK-QD 414
DB 353 TEVTRLASADDAAPGSPNSHVYQLSPBEBQVEBRAQVPTSGSVTLGVPLRAGN 412
QY 415 TPQVNLTVESDMD--FKTLCPQVNIINDIPIFEKSDVGNLTLAEPNIGSTILTI 472
DB 413 ILLVILAMDLAGEGFSSTCEVEAVTINDHAPFITSQIGPISLPEDVEGTIVML 472
QY 473 QATDAD-BEPTGSSKIL-YHIKGSBGRIGVDTDPHTNTGVYIK--KPLDETAASN 528
DB 473 TAIADALEP---AFRLMDFAIERGDTGTFGLDMEB--DSGHVRLRLCKNLSEEAAPSE 527
QY 529 IYKAKENPEPLVFGVKYNASFPKFTLYTDVNEAQFGQHVFOQAYSEDAVIGTQVNV 588
DB 528 VVVVQSVAKLV-GPGPGGATATVTLVERVWPPKLOESYEASVPIAPAGSFLTLTI 586
QY 589 TAKDPREGLDISYSLRGDTGRWLKIDHVTGEIFSAVPLD-REAGSPRYOVVATEVGSGL 647
DB 587 QPSDPISTRFSLVNDSSBWLCTEKFSGEHTAQSLOQAQPDITYVLEAODTA---- 642
QY 648 SSVSEFHLILMDVNDNPRLAKDYGLFFCHPLSAPSGSLF-BATDDOHLFRGPHFTFS 706
DB 643 -----LTLAPVPSQ-----YLCPTPQDHGLVSGPSKPDLASGCP-YSFT 683
QY 707 LG-SGSLQNDWEVSKINGTHARLSTHTDPEEPAVYVLRINDGSRPPLGLYSLEPTPC 765
DB 684 LGBNPTVORDMRLOQTINGSHAYVTLALHWVEPREHIIPVVSHNAQ---MMQLLVAYIVC 740
QY 766 SC-VEGSCPRPAGHQIGPTVGAAGVILTLTLVIGIILAVVIRIKDKGKNVSACA 824
DB 741 RCNVEGQCKRKVKRMKMPKLSAVGLVGTILVAILFILITHTMTMSKKDPDPGADS 800
QY 825 SEVK 828
DB 801 VPLK 804

```

RESULT 12

US-10-006-856A-229

Sequence 229, Application US/10006856A

Publication No. US20030044841A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Baton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gueney, Austin L.
 APPLICANT: Hillan, Kenneth J.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 TITLE OF INVENTION: Secured and Transmembrane Polypeptides and Nucleic
 FILE REFERENCE: P2830P1C14
 CURRENT APPLICATION NUMBER: US/10/006,856A
 NUMBER OF SEQ ID NOS: 477
 Prior Application removed - See File Wrapper or Palm

SEQ ID NO 229
 LENGTH: 807
 TYPE: PR1
 ORGANISM: Homo sapiens
 US-10-006-856A-229

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

62 LTGE---TDNIFVIERE-GLVYNRALDRETRSTHNLQVLAALDANGIIVEGVPPIITYE 116
 57 LSGDSKATGEPFAMDPDGGFLLVTRALDREDAEQVQLVTEMDQGHVLMGQPVLYVH 116
 117 KDINDNRPFLQSKTEGVSVRONSRRGKPLVYNATDDPATPNQGLYYQIVQIOLPMINN 176
 117 KDENDQVPHFSQAIYRRLRSRGTGRIPLFLFLEASDDEGTANSDLRPHILISQAPOPS 176
 177 VMYFOINNTGALSLTRREGSOELNPAKNPSYLVISYKMGQSGSENSFSDTTSVDIYTE 236
 177 PMFQLEPRGLALALSPKSGTSLDHALERTYQLLVQVKMGQDA-SGHQATATVEAGIIE 235
 237 NIMKAPKPVEMVENSNDPHPIKITQVRMNDPGAQYSLVDEKULPRPFISIDQGDITYQ 236
 236 STWVSLERPHLAENLKVLYPHHMAQVMSGQDVHYHL---ESHPPGFVFNAGNLYVTR 232
 297 PLDREKDAYVFAVAKDEYGRPLSYPLFIHVKKDINDNPTCPSPVTVEVQENERLG 356
 293 ELDRQAQAEYLLQVRAONSHGEDYAPLELHVMDNDNVPICPRPDPVTSIPELSPPG 352
 357 NSIGTLTAHDEBENTANSEFLNRYIVQTPKLPMDG-LFLIQTAYAGMLQAKOSLKK-QD 414
 353 TEVTRLSAEDADAPGSPNSHVYQLLSPEDEGVEGRAQVDPPTSQSVTLGVPLRAGQN 412
 415 TPQYNLTIEVSDKD--FKTLCFVQIYNVIDINDQIPIFEKSDYGNLTAEPTNIGSTILTI 472
 413 ILLVLVAMDLAGAEGGFSSTCEVEVAVTINDNAPEFITSQIGPISLPEBVEGGLVAML 472
 473 QATDAD-EPTTSSKIL-YHIIKDGSEGRGLVDTPHTNTGYVLIK-KPLDEPETAASN 528
 473 TALDADIEP---AFRLMDFAIERGDTYGTGLDMEP--DSGHVRLRLCKNLSTYEAAPSH 527
 529 IYKAEINPEPLVEGVKYNASSFAKFTLIYTDVNEARPOFSGHVEQAKVSEDAITGVKNV 588
 528 VVVVVGVSQAKLV-GPGGPGATATVTVLVERVMPRLDDESEVSAIPAPAGSFLITI 586
 589 TAQDPGLDISYSLRGDTGMLKIDHVTGEIFSVAPLD-REAGSPRYOVVATEVGGSSL 647
 587 QPSPDPSRTRLFSLVNDSEGMLCIEKFSGEVHTAQSIOGAQPGDITYVLEAQDTA---- 642
 648 SSVSEPHLLMDVNDNPRRLAKDYTGLFCHPLSAPGSLF-BATDDQHLFRGPHFTS 706
 643 -----LTLAPPSQ-----YLCPTRODGLVIGSGSKDPDLASGHGP-YSFT 663
 707 LG-SGSLDNMEVSKINGTHARLSTRHTDPEERAYVLLIRINDGSRPLEGIVSLPVTFC 765
 684 LGNPNFYQRMRLQTLNGSHAYLTLALHWEPREHIIIPVYVSHNAQ---MMQLLVAYIVC 740
 766 SC-VESGCPRPAGHQGIFTVGNAVGLITLLTVIGIILAVNFRIRIKDQKGNVSAQA 824
 741 RCNVEGQCKRKVGRMKGMPTKLSAVGILVGLVLAIGIFLITLFTHTWMSHKDPPDADS 800
 QY 825 SEVK 828
 DB 801 VPLK 804

RESULT 13
 US-10-218-631-188
 Sequence 188, Application US/10218631
 Publication No. US20030045687A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.

APPLICANT: Desnoyers, Luc
 APPLICANT: Geriltsen, Mary
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Guiney, Austin L.
 APPLICANT: Smith, Victoria
 APPLICANT: Stephan, Jean-Philippe F.
 APPLICANT: Watanabe, Colin L.
 APPLICANT: Wood, William I.
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 TITLE OF INVENTION: ACIDS ENCODING THE SAME
 FILE REFERENCE: P3530P1C14
 CURRENT APPLICATION NUMBER: US/10/218,631
 PRIOR FILING DATE: 2002-08-12
 PRIOR APPLICATION NUMBER: 10/119,480
 PRIOR FILING DATE: 2002-04-09
 PRIOR APPLICATION NUMBER: 60/059113
 PRIOR FILING DATE: 1997-09-17
 PRIOR APPLICATION NUMBER: 60/062287
 PRIOR FILING DATE: 1997-10-17
 PRIOR APPLICATION NUMBER: 60/063549
 PRIOR FILING DATE: 1997-10-28
 PRIOR APPLICATION NUMBER: 60/064103
 PRIOR FILING DATE: 1997-10-31
 PRIOR APPLICATION NUMBER: 60/069873
 PRIOR FILING DATE: 1997-12-17
 PRIOR APPLICATION NUMBER: 60/078910
 PRIOR FILING DATE: 1998-03-20
 PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079656
 PRIOR FILING DATE: 1998-03-26
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 Remaining Prior Application data removed - See file wrapper or PAM.
 NUMBER OF SEQ ID NOS: 246
 SEQ ID NO 188
 LENGTH: 807
 TYPE: PR1
 ORGANISM: Homo Sapien
 US-10-218-631-188

Query Match 21.0%; Score 908.5; DB 9; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.4e-59;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

62 LTGE---TDNIFVIERE-GLVYNRALDRETRSTHNLQVLAALDANGIIVEGVPPIITYE 116
 57 LSGDSKATGEPFAMDPDGGFLLVTRALDREDAEQVQLVTEMDQGHVLMGQPVLYVH 116
 117 KDINDNRPFLQSKTEGVSVRONSRRGKPLVYNATDDPATPNQGLYYQIVQIOLPMINN 176
 117 KDENDQVPHFSQAIYRRLRSRGTGRIPLFLFLEASDDEGTANSDLRPHILISQAPOPS 176
 177 VMYFOINNTGALSLTRREGSOELNPAKNPSYLVISYKMGQSGSENSFSDTTSVDIYTE 236
 177 PMFQLEPRGLALALSPKSGTSLDHALERTYQLLVQVKMGQDA-SGHQATATVEAGIIE 235
 237 NIMKAPKPVEMVENSNDPHPIKITQVRMNDPGAQYSLVDEKULPRPFISIDQGDITYQ 296
 236 STWVSLERPHLAENLKVLYPHHMAQVMSGQDVHYHL---ESHPPGFVFNAGNLYVTR 292
 297 PLDREKDAYVFAVAKDEYGRPLSYPLFIHVKKDINDNPTCPSPVTVEVQENERLG 356
 293 ELDRQAQAEYLLQVRAONSHGEDYAPLELHVMDNDNVPICPRPDPVTSIPELSPPG 352
 357 NSIGTLTAHDEBENTANSEFLNRYIVQTPKLPMDG-LFLIQTAYAGMLQAKOSLKK-QD 414
 353 TEVTRLSAEDADAPGSPNSHVYQLLSPEDEGVEGRAQVDPPTSQSVTLGVPLRAGQN 412
 415 TPQYNLTIEVSDKD--FKTLCFVQIYNVIDINDQIPIFEKSDYGNLTAEPTNIGSTILTI 472

```

Db      413 ILLVLAMDLAGEGFSSTCEVEAVTIDNDHAPFITSQIGPISLPEDEPGTLVAML 472
Qy      473 QATDAD-EPTGSSKLL-YHIIKDSGRLGVDTDPHTNTGYIIR--KPLDETAASN 528
Db      473 TAIADALEP---AFRLMDFAIERGDTGTFGLDMEP--DSGHVRLRLCKNLSTEAAPSHS 527
Qy      529 IVFKAENPEPLVGVKNASSPAKFTLLIVTDVNEAPQFSQHVQAKYSEDAVIGTKVNV 588
Db      528 VVVVVOGVAKLV-GPGPGCATVTVLVERVMPPLKDESEIASEAVPISAPAGSFLTTI 586
Qy      589 TAKDPGLDISYSLRGPTRGMLKIDHTGEIFSVAPLD-REAGSPYRVQVATEVGGSSL 647
Db      587 QPDPISRTLRFLVNDSEGLCEKFSGEVHTAQSLQGAQPGDVTYVLEAODTA---- 642
Qy      648 SSVSEFHLIMDVNDNPRKLAKDYTGILFCHPLSAPGSLIF-EATDDQHLFRGPHRTFS 706
Db      643 -----LTLAPVPSQ-----YLCPTRODHGLIVSGPSKDPDLASGHGP-VSFT 683
Qy      707 LG-SGSLONDMVSKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLGIVSLPYTFC 765
Db      684 LGPPTVQDRMRLQTLNGSHAYLTLALHWEPREHIIIPVVSHNAQ---MMQLLVRIYVC 740
Qy      766 SC-VEGSCFRPAGHQTGIPVGNAGVILTTLLVIGIILAUVFIRIKDKGKNVESQA 824
Db      741 RCNVEGQCRRKVGKMGMPKLSAVGILVGTVAIGIFILLIFTHMTMSRKDPDQPADS 800
Qy      825 SEVK 828
Db      801 VPLK 804

RESULT 14
US-10-230-338-188
; Sequence 188, Application US/10230338
; Publication No. US20030044934A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerdtisen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P350P1C92
; CURRENT APPLICATION NUMBER: US/10/230.338
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

```

; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 188
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-230-338-188

Query Match      21.0%; Score 908.5; DB 9; Length 807;
Best Local: Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;

Qy      62 LTGE-----TDNIFVIERE-GLLYNRALDRETRSTNHLQVAIDANGIIEGVPITIBV 116
Db      57 LSGDSKATGEGPAMPDPSGFLVTRALDREBOAEYQLQVLEMOGHVLMGQPYLVHY 116
Qy      117 KDINDRPFPLQSKYSGSRQNSRPEKPFILYNNATDLDPRPNGLYQIVQLQMINN 176
Db      117 KQENDQVPHRSQAIYARLSRGRPGIPPLFLASRDRPPTANSDLRPHILISQAPADS 176
Qy      177 VMVFOINNTGAIISLRREGSOELNPAKNSYMLNIVSKMGQSENSFSDTTSVDIITYE 236
Db      177 PMFQLEPRGLALALSPKSTSLDHALERTYQLLVQYKMGQDA-SGHQATINVEISIE 235
Qy      237 NIWKAEPVEMWENSTDPPIKITQYRMNDPEAQYSLVDKELPRFPFSIDQEGDIYVTO 296
Db      236 STWVSLPFIHLENTLVLYPHHMAQVHMGSDVHYHL---ESHPPGFVEVNAEGNLVYTR 292
Qy      297 PLDRKDAVVFVAVAKDEYKPLSYPLFIHYVKQINDNPPCPSPVTVFEQENBRIG 356
Db      293 ELDRQAQAEYLLQVRAQNSHGEDYAPLEHLVLMDENNVPCPRPDPTVSIPELSPPG 352
Qy      357 NSIGTLTAHDREENTANSPFNRIYEQTPKLPMDG-LFLICTYAGMLQAKOSLKK-OD 414
Db      353 TEVTRLSAEDADAPGSPNSHVYQLSPREBQVEGRAQVDPDTSSVTLGVPLRAQGN 412
Qy      415 TPQYNLTIEVSDKD-FKTLCPVQINVIDINDQIPFEKSDVGNLTIAEDTNGITLTI 472
Db      413 ILLVLAMDLAGEGFSSTCEVEAVTIDNDHAPFITSQIGPISLPEDEPGTLVAML 472
Qy      473 QATDAD-EPTGSSKLL-YHIIKDSGRLGVDTDPHTNTGYIIR--KPLDETAASN 528
Db      473 TAIADALEP---AFRLMDFAIERGDTGTFGLDMEP--DSGHVRLRLCKNLSTEAAPSHS 527
Qy      529 IVFKAENPEPLVGVKNASSPAKFTLLIVTDVNEAPQFSQHVQAKYSEDAVIGTKVNV 588
Db      528 VVVVVOGVAKLV-GPGPGCATVTVLVERVMPPLKDESEIASEAVPISAPAGSFLTTI 586
Qy      589 TAKDPGLDISYSLRGPTRGMLKIDHTGEIFSVAPLD-REAGSPYRVQVATEVGGSSL 647
Db      587 QPDPISRTLRFLVNDSEGLCEKFSGEVHTAQSLQGAQPGDVTYVLEAODTA---- 642
Qy      648 SSVSEFHLIMDVNDNPRKLAKDYTGILFCHPLSAPGSLIF-EATDDQHLFRGPHRTFS 706
Db      643 -----LTLAPVPSQ-----YLCPTRODHGLIVSGPSKDPDLASGHGP-VSFT 683
Qy      707 LG-SGSLONDMVSKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLGIVSLPYTFC 765
Db      684 LGPPTVQDRMRLQTLNGSHAYLTLALHWEPREHIIIPVVSHNAQ---MMQLLVRIYVC 740
Qy      766 SC-VEGSCFRPAGHQTGIPVGNAGVILTTLLVIGIILAUVFIRIKDKGKNVESQA 824
Db      741 RCNVEGQCRRKVGKMGMPKLSAVGILVGTVAIGIFILLIFTHMTMSRKDPDQPADS 800
Qy      825 SEVK 828
Db      801 VPLK 804

RESULT 15
US-10-063-518-98
; Sequence 98, Application US/10063518
; Publication No. US200300449735A1
; GENERAL INFORMATION:

```


APPLICANT: Eaton, Dan L.
APPLICANT: Falvarcoff, Ellen
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Goddard, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gueney, Austin L.
APPLICANT: Watanabe, Colin K.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3230R1C1
CURRENT APPLICATION NUMBER: US/10/063,518
PRIORITY FILING DATE: 2002-05-01
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 98
LENGTH: 807
TYPE: PRT
ORGANISM: Homo Sapien
US-10-063-518-98

Query Match 21.0%; Score 908.5; DB 9; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.4e-59;
Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
QY 62 LTGE---TDNIFIERE-GILYRNALDRETRSTHNLQYAAALDANGIIVEGPVITIEV 116
DB 57 LSGDSGKATGPRFMDSDGFLVTRALDREEQAYQLQVTLQEMQGHVLMGPOPLVHV 116
QY 117 KDINDNPTLQSKTEGSSVQRNSRPGKFLVYNATDLDDPATNGQLYQIVIQIPMINN 176
DB 117 KDEDDQVPHFSQAIYRRLSRGTGPIPLFLASDRDEPTANSDLRPHILSQAPQPS 176
QY 177 VWFQIINKKGALSTREGSQELNPAKPSNLYISKMDGGGSENSFSDTSYDIYTE 236
DB 177 PDMFQLEPRIGALALSPKSGTSLDHALERTYQLLVQKMDQDA-SGQATATVEYSIIE 235
QY 237 NIKAPKPEVMEVNSTDPHPKIKITQVWMDPGAQYSLVDEKELPRFPFSIDQEGDIYVTO 296
DB 236 STWVSEPIHIAENLKLTVHMAQVMSGDVHYHL---ESHDPGFVNAEGNLVYTR 292
QY 297 PLDREKDAYVFAVADKEXKPLSYPLKLVKVKDINDNPTCPSPVTFVEVOENERLG 356
DB 293 ELDRQAQAEYLLQVRAQNSHGEDYAAFLVLMENDNVPICPPRDPTVSIPELSPPG 352
QY 357 NSIGTLTAHNRDEENTANSFLNRIIVEQTRKLPMDG-LFLIQYAGMLQAKOSLKK-QD 414
DB 353 TEVTRLSAEDADAPGSPNSHVYQLSPBEDGVEGRAFOVDPTSGSVTLGVLPFRAGQN 412
QY 415 TPQYNLTIEVSDKD--FKTLCFQVQINVIDINDQPIPEKSDYGNLTLAEDTNIGSTILTI 472
DB 413 ILLVLAMDLAAGBGGFSTCEVEVAATVINDHAPETITSQIGPISLPEDVEPPTLVAML 472
QY 473 QATDAD-EPPTGSSKIL-YHIKGDSEGLGVDTPHTNTGYVLIK--KPLDPEFAVSN 528
DB 473 TALDADLEP--AFRLMDFAIERGDTGTFGLDWER--DSGHVRLRCKLSYEAPSHSHE 527
QY 529 IVKAEKPEPLVGVKYNASSFAKFTLIVTDVNEAPQFSQHVFOAKYSEDAVIGTKGVN 588
DB 528 VVVVVSVAKLV-GPGPGATATVTVLVERVMPPLDQESYEASVPIAPAGSFLTLTI 586
QY 589 TAKDPEGLDISYLRGTRGMLKIDHTGEIISVAPLD-REAGSPYRVQVATEVGGSSL 647
DB 587 QPSDPTISRTIRFSLVNDSEGLCTEKFSGEVHTAQSLQAGQDGTIVLVEADDTA---- 642
QY 648 SSVSEPHLIMDVNDNPPRLAKDYTGILFCHPLSAPGSLIF-EATDDQHLFRGPHFTFS 706
DB 643 -----LTLAPVPSQ-----YLCPTPRQDHGLIVSGPSKDPDLASGHGP--YSFT 683
QY 707 LG-SGSLQNMWESKINGTARLSTRITDPEERAYVVLIRINDGGRPPLGIVSLPPTFC 765
DB 684 LGNPPTVQDRMLOTINGSHAYLTLALHWVEPREHIIIPVVVSHNAQ---MWQLLVRYIVC 740

QY 766 SC-VEGSCFRPAGHQGTPTVGNVGIILTTLLVIGIILAVFIRIKDKGKNVESQA 824
DB 741 RCNVEGQCMKRVGRMKGMPTKLSAVGIVGTVAIGFILITHTMTMSRKKOPDPADS 800
QY 825 SEVK 828
DB 801 VPLK 804

Search completed: June 22, 2003, 01:10:50
Job time : 66 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:10:55 ; Search time 274 Seconds
(without alignments)
4455.837 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIQAMHSLCLMLYATG.....DKGKNVSAQASEVPLRS 832

Scoring table:
BLSDUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1042519 seqs, 73373590 residues
Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xid
-O=/cgn2_1/USPTO.spool/US10025380/runat_17062003_173329_24718/app query.fasta_1.1031
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmpb -MINMATCH=0.1
-LOOPEX=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blodun62
-TRANS=human40.csl -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10025380@cgn.1.1.166@runat_17062003_173329_24718
-NCPU=6 -ICPU=3 -NO MAP -LARGESQURRY -NEG SCORES=0 -WAIT -DSBLOC=100 -24718
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELext=7

Database : Published Applications NA:

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4321	100.0	3345	9	US-10-025-380-1076
2	4321	100.0	3345	10	US-09-922-217-1076
3	4321	100.0	3345	10	US-09-962-436-302
4	4321	100.0	3345	10	US-09-833-263-1076

5	4321	100.0	3345	10	US-09-880-107-3288	Sequence 3288, Ap
6	4303	99.6	3654	10	US-09-962-832-1118	Sequence 118, Ap
7	4303	99.6	3654	10	US-09-880-107-3807	Sequence 3807, Ap
8	4208	97.4	2877	9	US-10-025-380-1086	Sequence 1086, Ap
9	4208	97.4	2877	10	US-09-922-217-1086	Sequence 1086, Ap
10	4208	97.4	2877	10	US-09-833-263-1086	Sequence 1086, Ap
11	908.5	21.0	2848	9	US-10-063-547-97	Sequence 97, Appl
12	908.5	21.0	2848	9	US-10-063-516-97	Sequence 97, Appl
13	908.5	21.0	2848	9	US-10-063-502-97	Sequence 97, Appl
14	908.5	21.0	2848	9	US-10-227-884-187	Sequence 187, Appl
15	908.5	21.0	2848	9	US-10-230-163-187	Sequence 187, Appl
16	908.5	21.0	2848	9	US-10-006-556-228	Sequence 228, Appl
17	908.5	21.0	2848	9	US-10-218-631-187	Sequence 187, Appl
18	908.5	21.0	2848	9	US-10-230-338-187	Sequence 187, Appl
19	908.5	21.0	2848	9	US-10-063-118-97	Sequence 97, Appl
20	908.5	21.0	2848	9	US-10-063-598-97	Sequence 97, Appl
21	908.5	21.0	2848	9	US-10-227-693-97	Sequence 97, Appl
22	908.5	21.0	2848	9	US-10-230-614-187	Sequence 187, Appl
23	908.5	21.0	2848	9	US-10-006-818-228	Sequence 228, Appl
24	908.5	21.0	2848	9	US-10-015-393-228	Sequence 228, Appl
25	908.5	21.0	2848	9	US-10-063-567-97	Sequence 97, Appl
26	908.5	21.0	2848	9	US-10-216-159-187	Sequence 187, Appl
27	908.5	21.0	2848	9	US-09-946-374-228	Sequence 228, Appl
28	908.5	21.0	2848	9	US-10-012-121-228	Sequence 228, Appl
29	908.5	21.0	2848	9	US-10-015-869-228	Sequence 228, Appl
30	908.5	21.0	2848	9	US-10-218-849-187	Sequence 187, Appl
31	908.5	21.0	2848	9	US-10-227-873-187	Sequence 187, Appl
32	908.5	21.0	2848	9	US-10-227-883-187	Sequence 187, Appl
33	908.5	21.0	2848	9	US-10-063-599-97	Sequence 97, Appl
34	908.5	21.0	2848	9	US-10-219-076-187	Sequence 187, Appl
35	908.5	21.0	2848	9	US-10-230-434-187	Sequence 187, Appl
36	908.5	21.0	2848	9	US-10-006-116-228	Sequence 228, Appl
37	908.5	21.0	2848	9	US-10-017-527-228	Sequence 228, Appl
38	908.5	21.0	2848	9	US-10-013-913-228	Sequence 228, Appl
39	908.5	21.0	2848	9	US-10-017-527-228	Sequence 228, Appl
40	908.5	21.0	2848	9	US-10-063-595-97	Sequence 97, Appl
41	908.5	21.0	2848	9	US-10-219-075-187	Sequence 187, Appl
42	908.5	21.0	2848	9	US-10-219-075-187	Sequence 187, Appl
43	908.5	21.0	2848	9	US-10-219-466-187	Sequence 187, Appl
44	908.5	21.0	2848	9	US-10-219-466-187	Sequence 187, Appl
45	908.5	21.0	2848	9	US-10-219-479-187	Sequence 187, Appl

ALIGNMENTS

RESULT 1
US-10-025-380-1076
Sequence 1076, Application US/10025380
Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodges, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Yuguang
APPLICANT: Jiang, Yuguang
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yaeir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1076
 LENGTH: 3345
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-025-380-1076

Alignment Scores:

Score:	4321.00	Length:	3345
Percent Similarity:	100.00%	Matches:	832
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	100.00%	Mismatches:	0
		Indels:	0
		Gaps:	0

US-10-025-380-1081 (1-832) x US-10-025-380-1076 (1-3345)

Qy 1 MetIleuGlnAlaHisIleuHisSerIeuCysIeuIeuMetIeuTYrIeuAlaThrGly 20
 Db 88 ATGATACCTTGAAGCCCATCTTCACTCCCTGTGTCTTCTTATGTCCTTATTTGGCAACTGGA 147
 Qy 21 TyGIgInGInGlyIyIyPheSerGlyProIeuIyPheMetThrPheSerIleTYrGlu 40
 Db 148 TATGCGCAAGAGGGGAAAGTTAGTGGAACCCCTGAAACCCATGACATTTTCTATTATGAA 207
 Qy 41 GIVGInGInPProSerGInIleIlePheGInPheIySaIaenProProlaValThrPhe 60
 Db 208 GGGCAAGAACCGAGTCATATTATTTCCAGTTTAAAGCCCATCTCTCTGTGACTTTT 267
 Qy 61 GluIeuThrGlyGInThrAspAsnIlePheValIleGluArgGluGlyIeuIeuTYr 80
 Db 268 GAACTTACTGGGGAGAGACAGACATATTTGTGATAGACGGGAGGACTTCTGTATTAC 327
 Qy 81 AsnArgAlaIeuAspArgIuThrArgSerThrHisasnIeuGInValAlaIleuAsp 100
 Db 328 AACGAGCGCTTGACAGGAAACAAGATCTACTCAATCTCCAGTTTGCAGCCCTGAC 387
 Qy 101 AlaasnGlyIleIleValGluGlyProValProIleThrIleGluValIyAspIleAsn 120
 Db 388 GCTATGGAATTTATAGTGGAGGGTCCAGTCCCTTACCATGAAAGTGAAGACATCAAC 447
 Qy 121 AspAsnArgProThrPheIeuGInSerIyTYrGluIySerValArgGInAsnSerArg 140
 Db 448 GACATGACCCCACTGTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAAGGACATCTGCG 507
 Qy 141 ProGlyIyPProIleIeuTYrValIAsnAlaThrAspIeuAspProAlaThrProAsn 160
 Db 508 CCAGAAAGCCCTTCTGTATGATGCATGCCAGACCTGGATGATCCGGCCACTCCCAT 567
 Qy 161 GIVGInIeuTYrTYrGInIleValIleGInIeuPProMetIleAsnAsnValMetTYrPhe 180
 Db 568 GGGCAGCTTTATTCAGATTTGTCTCCAGCTTCCCATGATCAACATGTCAATGTACTTT 627
 Qy 181 GInIleAsnAsnIyThrGlyAlaIleSerIeuThrArgGluIySerGInIuIeuAsn 200
 Db 628 CAGATCAACAACAAGGAGGAGGACATCTCTTACCCGAGAGGATCTCAGGAATTGAT 687
 Qy 201 ProAlaIyAsnProSerTYrAsnIeuValIleSerValIyAspMetGlyIyGInSer 220
 Db 688 CTTGCTTAAGATCTTCTTATATCTGTGATCTCAGTGAAGGACATGGGAGGCGCAGGT 747
 Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValITrGluAsnIleTrpIyS 240
 Db 748 GAGATTCCTTCAGATCAACATCTGTGATATCATATGACAGAGAAATATTTGGAA 807
 Qy 241 AlaProIyAspProValGluMetValGluAsnSerThrAspProHisProIleIyS 260
 Db 808 GCACCAAAACCTGTGGAGATGTTGGAATACTCACTGATCTCAACCCCATCAAAATCACT 867
 Qy 261 GInValIArgTrpAsnAspProGlyAlaGInTYrSerIeuValAspIyGluIyIeuPro 280
 Db 868 CAGGTGGCGGTGAATGATCCGGTGCACAATATTTCTTATGTTGACAAAGAAAGACTGCA 927
 Qy 281 ArgPheProPheSerIleAspGInGluIyAspIleTYrValIThrGInProIeuAspArg 300

Db 928 AGATCCCATTTTCAATTGACAGAGAGATATTACGTGACTCAGCCCTTGACCGCA 987
 Qy 301 GluGluIyAspAlaTYrValPheTYrAlaValAlaIyAspGluTYrGlyIySProIeu 320
 Db 988 GAAAGAAAGATGCAATATGTTTTTTATGCAAGTGGCAAGAGATGATACGAAACCACTT 1047
 Qy 321 SerTYrProIeuGluIleHisValIyValIyAspIleAsnAspAsnProProThrCys 340
 Db 1048 TCATATCCCTGGAATTTCTATGTAAGTTAAAGATTTATATATATATCACTTACATG 1107
 Qy 341 ProSerProValIyThrValPheGluValGInIuAsnGluArgIeuGlyIyAsnSerIleGly 360
 Db 1108 CCGTCACAGTAAACCGATATTGAGTCCAGAGATGAACGACTGGGTAAACATATCGGG 1167
 Qy 361 ThrIeuThrAlaHisAspArgAspGluIuAsnThrAlaAsnSerPheIeuAsnTYrArg 380
 Db 1168 ACCCTTACGTGACATATACAGGATGAAGAAATACGCAACGTTTCTTAACTACAGG 1227
 Qy 381 IleValGluGInThrProIySleuPProMetAspGlyIeuPheIeuIleGInThrTYrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAACTTCCCATGATGAGACTCTTCTTATCCAACTATGCT 1287
 Qy 401 GlyMetIeuGInIeuAlaIySerIeuIyValyGInAspThrProGInTYrAsnIeu 420
 Db 1288 GAATGTACAGTTACTTAAACAGTCTTGAAAGAACAGATACTCTCAGTACAACTTA 1347
 Qy 421 ThrIleGluValSerAspIyAspPheIyThrIeuCysPheValGInIleAsnValIle 440
 Db 1348 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTGTGCAAACTGTTATT 1407
 Qy 441 AspIleAsnAspGInIleProIlePheGluIySerAspTYrGlyAsnIeuThrIeuAla 460
 Db 1408 GATATCATATGATAGATCCCATCTTTGAAAATACGATTTATGAAACCTGACTTGTCT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleIeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAGACAAACATTTGGGTCCACCATCTTAAACCATCAGGCCACTGATGAGACCA 1527
 Qy 481 PheThrGlySerSerIySleuTYrHisIleIleIyGlyAspSerGluIyArgIeu 500
 Db 1528 TTTTACCTGGAGTCTTAAATTTCTGTATCATATATGAAGGAACAGTGGAGGAGCTTG 1587
 Qy 501 GIVAlaAspThrAspProHisThrAsnThrGlyTYrValIleIySlySProIeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATCAACACCGGATATGTCTATATTAAGGCTCTTGAT 1647
 Qy 521 PheGluThrAlaAlaIySerAsnIleValPheIyAlaGluAsnProGluProIeuVal 540
 Db 1648 TTTGAAACAGCAGCTGTTTCAACATTTGTGTTCAAGCAGAAATCTCGAGCTTATGTG 1707
 Qy 541 PheGlyValIyTYrAsnAlaSerSerPheAlaIyPheThrIeuIleValIThrAspVal 560
 Db 1708 TTTGTGTGAAAGTACATGCAAGTCTTTTGGCAAGTTCACGCTTATTTGTGACAGATGTG 1767
 Qy 561 AsnGluAlaAspProIlePheSerGInIleValPheGInAlaIyValSerGluAspValAla 580
 Db 1768 AATGAAACACTCAATTTTCCCAACAGTATTTCCAGCGAAATCACTGATGATGATGCT 1827
 Qy 1828 ATAGGCACTTAAATGTGGCAATGTGATGCCAAGATCCAGAAAGTCTGACATATGCTAT 1887
 Db 601 SerIeuArgIyAspThrArgGlyTYrPheIyIleAspHisValIThrGlyIyGluIlePhe 620
 Db 1888 TCACTGAGGGAGACACAAAGGTTGGCTTAAATTTGACCACTGATCTGTGTGATCTTT 1947
 Qy 621 SerValAlaProIeuAspArgIuIyAGIySerProTYrArgValGInValAlaThr 640
 Db 1948 AGGTGGCTCATTTGGACAGAAAGCCGGAAGTCCATATGTGGTACAAAGTGGTGGCCACA 2007
 Qy 641 GluValGlyIySerSerIeuSerValSerGluPheHisIeuIleIeuMetAspVal 660

Db 2008 GAAGTACGGGGCTCTTCTTAAGCTGTGTCTGAGAGTCCAGCTGATCCTGATCCTTAATGATGTG 2067
 QY AsnAspAsnProProAlaGluAlaAlaAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCACGAGCTAGCCAAAGACTACACGGGCTGTGTTCTTCTCCATCCCTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2128 AGGACACCTGGAAGTCTCATTTTCGAGCTACGATGATGATCAGCACTTATTTCGGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLeu 720
 Db 2188 CCCATTATTAATTTTCCCTCGGAGAGGAGCTTACAAAACAGCTGGGAAGTTTCCAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaArgAlaTyr 740
 Db 2248 ATCAATGATCTCATGCGGCACTGTCTACAGGACACACACTTTGAGGAGGGCGTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 Db 2308 GTGCTCTTGATCCGCATCAATGATGGGGGTGCGCCACCTTGGAAGGCAATTGTTCTTGA 2367
 QY 761 ProValThrPheCysSerCysValGlyGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTACATTCGACAGTTGTGTGGAAGAGTTGTTTCCGGCAGCAGGTCACAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCCACTGTGGGCACTGCGATGGTGTGATCTGTGACACCTTCTGCGATTTGAT 2487
 QY 801 IleIleLeuAlaValAlaPheIleArgIleValLysAspLysGlyLysAspAsnValGlu 820
 Db 2488 ATAAATTTAGCAGTTGTGTTATCCGCATTAAGAGATTAAGCAAAAGATTAATTTGAA 2547
 QY 821 SerAlaGluAlaSerGluValLysProLeuArgSer 832
 Db 2548 AGTCTCAAGCATCTGAAGTCAAACTCTGAGAAGC 2583

RESULT 2

US-09-922-217-1076

Sequence 1076, Application US/09922217

Patent No. US2002076414A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Lodes, Michael J.

APPLICANT: Secretist, Heather

APPLICANT: Benson, Darin R.

APPLICANT: Meagher, Madeleine Joy

APPLICANT: Stolk, John A.

APPLICANT: Wang, Tonglong

APPLICANT: Jiang, Yugu

APPLICANT: Smith, Carole Lynn

APPLICANT: King, Gordon E.

APPLICANT: Wang, Aijun

APPLICANT: Clapper, Jonathan D.

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

FILE REFERENCE: 210121.471C13

CURRENT APPLICATION NUMBER: US/09/922.217

NUMBER OF SEQ ID NOS: 1124

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1076

LENGTH: 3345

TYPE: DNA

ORGANISM: Homo sapiens

US-09-922-217-1076

Alignment Scores:

Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-025-380-1081 (1-832) x US-09-922-217-1076 (1-3345)

QY 1 MetIleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 88 ATGATCTTCAGGCGCCATCTTCACTCCCTGTGTCTTATATGCTTATTTGCACTGCA 147
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 Db 148 TATGGCCAAAGGGGAAGTTTATGTGATCCCTGAAAACCAATGACATTTCTATTTAGAA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPhe 60
 Db 208 GGCCAAGAACCAAGTCAATTAATTCAGATTAAAGCCCAATCTCTGCTGAGACTTT 267
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80
 Db 268 GAACATACTGGGGAGACAGACAAACATATTTGTGATGAACGGGAGGAGACTTGTATTTAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 328 AACAGAGCCTTGACAGAGGAACAAAGATCTACTACAAATCTCCAGGTTGACAGCCCTGAC 387
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db 388 GCTAATGAAATTAATGATGAGGGTCCAGTCCCTATACCAATGAATGAAGACATACAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg 140
 Db 448 GACAAATCGAACCAAGTTTCTCCAGTCAAGTACGAAGGCTCAGTAAGCGAACTCTCCG 507
 QY 141 ProGlyLysProPheLeuTyrValAlaAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 508 CCAGGAAGCCCTTCTGTATGATCAATGACACAGACTGATGATCCGCGCACTCCCAAT 567
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGCACGCTTTATTAACAGATGTGTCACTCAGCTTCCATATCAACATGATGATGACTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 628 CAGATCAACAAACAAACGGGAGCATCTCTTTACCCGAGAGGAGCTCAGGAATGGAAT 687
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
 Db 688 CCGCTTAAGAAATCCCTTCTATATCTGATCTCAGTGAAGACATGGAGGCCAGAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
 Db 748 GAGAAATCTTTGATGATATCCACATCTGTGATATCATATGACAGAAATATTTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 808 GACCAAAACCTGTGAGATGTGAGAAACTCACTGATCCCAACCCCATCAAAATCACT 867
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 868 CAGTCCGGGTGAATGATCCCGTGACAAATTTCTTGTGTGACAAAGAAAGTGGCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
 Db 928 AGATTCCTCATTTTCATATGACAGAGAGAAATTTACGTACCTAGCCCTTGACCA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaAlaAspGluTyrGlyLysProLeu 320
 Db 988 GAAGAAAGAGATGATATGTTTTTATGCAAGTTGCAAGAGATGATGAGAAACCACTT 1047
 QY 321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 Db 1048 TCATATCCCGCTGAAATTCATATGAATTAAGATATTAAGATATTCACCTACATGT 1107

QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB 1108 CCGTCACCACTAACCGGATTTGAGTCCAGAGAAATACGATCGGTAAACAGATCGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 DB 1168 ACCCTACTCACAATGACAGGATGAGAAATACTGCCAAAGTTTCTTAACTACAG 1227
 QY 381 IleValGlnGlnThrProGlyLeuPheProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 DB 1228 ATGTGGAGCAAACTCCCAAACTTCCATGATGAGATCTTCTTAAATCCAAACTATGCT 1287
 QY 401 GluMetLeuGlnLeuAlaValGlnSerLeuGlyGlnAspThrProGlnTyrAsnLeu 420
 DB 1288 GGAATGTTACAGTTAGCTTAAACAGTCTTGAAGAGCAAGTACTCTCTAGTACACTTA 1347
 QY 421 ThrIleGluValSerAspGlyAspPheLeuThrLeuGlySerPheValGlnIleAsnValIle 440
 DB 1348 ACGATAGAGGTGTGCAAAAGATTTCAGACCTTTGTTTGTGCAAAATCAAGTTATT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGlnIleSerAspTyrGlyAsnLeuThrLeuAla 460
 DB 1408 GATATCATATATATGATCCCATCTTGAATAATCAGATTATGAAACCTGACTTGTCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAGACCAAAACATTTGGGTCCACATCTTAAACCTCCAGGCCACTGATGATGACCA 1527
 QY 481 PheThrGlySerSerIleLeuTyrHisIleIleLeuGlyAspSerGlnGlyArgLeu 500
 DB 1528 TTTACTGGAGTCTTAATAATCTGTATCATCATTAAGGAGACAGTGAAGGAGCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAspThrGlyTyrValIleIleLeuGlySerProLeuAsp 520
 DB 1588 GGGGTTCACACAGATCCCATCCCAACACCGGATATGTCATTAATAAAGCCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheValGluAsnProGluProLeuVal 540
 DB 1648 TTTGAAACAGACGTGTTTCCACATCTGTGTTCAACAGAAATCTGAGCCCTGATG 1707
 QY 541 PheGlyValIleTyrAsnAlaSerSerPheAlaValSerPheThrLeuIleValThrAspVal 560
 DB 1708 TTTGGTGTGAAGTACATGCAAGTTCTTTTCCAACTTCCAGCTTATTTGACAGATGTG 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaValSerGluAspValAla 580
 DB 1768 AATGACACACTCAATTTTCCCAACAGATATCCAAACGAAAGTCAGTGAAGATGTGCT 1827
 QY 581 IleGlyThrIleValGlyAsnValThrAlaValAspProGluGlyLeuAspIleSerTyr 600
 DB 1828 ATAGGCACTAAGTGGGCAATGTATCTGCCAAGATCCAAAGGTCTGGACATTAAGCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLeuValIleAspHisValIleThrGlyValIlePhe 620
 DB 1888 TCACGTGAGGGGACACACAGAGGTGGCTTAAATTAATGACCACTGAGTGGAGATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 DB 1948 AGGTGTGCTCCATTTGACAGAGGAGCGGAGTCCATATCGGTACAGATGTGTGCACA 2007
 QY 641 GluValGlyGlySerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspVal 660
 DB 2008 GAAATAGGGGGGTCTTCTTAAGCTCTGTGTCAAGTTCCACTGATCCCTTATGATGTG 2067
 QY 661 AsnAspAsnProProAlaGluValAlaValAspTyrThrGlyLeuPhePheCysHisProLeu 680
 DB 2068 AATGACAACTCCAGGCTAGCCAGACGATACAGGGCTTCTTCTGCAATCCCTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspArgGlnHisLeuPheArgGly 700
 DB 2128 AGTGCACCTGGAATCTCATTTTTCAGAGCTACTATATATGATGATGACCTTAATTTCCGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerTyr 720

DB 2188 CCCCATTTTACATTTTCCCTCGCAGTGAAGACTTACAAAGACTGGGAAAGTTTCCAAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluArgAlaTyr 740
 DB 2248 ATCAATGTATCTCAATGACCCGACGTGTACAGGACACAGACTTGGAGAGGGGCTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyIleArgProProLeuGlnGluIleValSerLeu 760
 DB 2308 GTGCTTTATCCGATCATATGATGAGGGGTCCGCCACCTTGGAAAGCATTTGTTCTTAA 2367
 QY 761 ProValThrPheCysSerCysValGlnGlySerCysPheArgProAlaGlyHisGlnThr 780
 DB 2368 CCAATTACATTTCTGACAGTTGTGTGGAGAAAGATTGTTCCGGCCAGACAGTCCACAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValIleLeuLeuThrThrLeuLeuValIleGly 800
 DB 2428 GGGATACCACTGTCGTGGCATGGCAGTGTGTATATCTGACACACCTTCTGTGATGGT 2487
 QY 801 IleIleLeuAlaValAlaPheIleArgIleValAspGlyValAspAsnValGlu 820
 DB 2488 ATATTTTACAGCTGTGTGTTATCCCATTAAGAAAGATTAAGGCAAAAGTATATGTGA 2547
 QY 821 SerAlaGlnAlaSerGluValIleValSerProLeuArgSer 832
 DB 2548 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAGAC 2583

RESULT 3
 US-09-962-436-302
 ; Sequence 302, Application US/09962436
 ; Patent No. US20020081301A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Soppel, Daniel
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
 ; FILE REFERENCE: 689290-75
 ; CURRENT APPLICATION NUMBER: US/09/962,436
 ; CURRENT FILING DATE: 2001-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235,082
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/234,924
 ; PRIOR FILING DATE: 2000-09-25
 ; NUMBER OF SEQ ID NOS: 568
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 302
 ; LENGTH: 3345
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-962-436-302

Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x US-09-962-436-302 (1-3345)

QY 1 MetIleLeuGlnAlaHisIleAsnSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 DB 88 ATATATCTTACAGCCCATCTTCACTCCCTGTGTCTTATGCTTATTTGCAACTGGA 147
 QY 21 TyrGlyGlnGluGlyIlePheSerGlyProLeuValProMetThrPheSerIleTyrGlu 40
 DB 148 TATGGCCAAAGAGGGGAAGTTTATGATGACCCCTGAAACCAATGACATTTTCTATTATGAA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAspProProAlaValThrPhe 60
 DB 208 GGCACAAAGACCGAGTCAATATATATTCAGTTTAAGGCCAATCTCTGCTGACCTTTT 267
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlnGlyLeuLeuTyrTyr 80

Db 268 GAACCTAATCGGGGAGACACACATATTTGTGATAGAACGGGGGAGCTTCTGATATAC 327
Qy AsnArgAlaLeuAspArgGluThrArgSerThrIshAsnLeuGluValAlaAlaLeuAsp 100
Db 328 AACGAGCCTTGGACAGGAGAACAGATCTTACTACATCTCCAGGTGGACCCCTGGAC 387
Qy AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIysAspIleAsn 120
Db 388 GCTATAGGAATTAATAGGAGAGGCTCCAGTCCCTATACCAATAGAGAGGAGCATCAAC 447
Qy AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValIaGlnAsnSerArg 140
Db 448 GACATATGACCCACGTTCTCCACTCAAGTACGAAAGCTCAAGAGGAGACTCTGCG 507
Qy ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db 508 CCGAGAAAGCCCTCTGTATGTCAATGCGACAGACTGGATGATCCGGCCACTCCCAAT 567
Qy GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db 568 GGCACGCTTTATACAGATTTGTCTACAGCTTCCATGATCAACATGTCTATCTT 627
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db 628 CAGATCAACAAACAAACGAGGACCATCTCTTACCCGAGAGGGATCTCAGGAATGAAAT 687
Qy ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
Db 688 CCGCTAGAAATCTTCTCTAATCTGTATCTGTATCTCAATGAAAGCAATGGAGGCAAGT 747
Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
Db 748 GAGATTCCTTCAAGTATCCACATCTGTGATATCTAGTACAGAGATATTTGGAAA 807
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProIshAspIleThr 260
Db 808 GCACCAAAACCTGGAGATGTGTGAAAACTCAACATGATCTCCACCAATCAAAATCACT 867
Qy 261 GluValArgThrAsnAspProGlyValGlnTyrSerLeuValAspLysGluLysLeuPro 280
Db 868 CAGGTGGGTGAGATGATCCCGGTGACAAATATCTTACTGTGCAAAAGAGAACTGCCA 927
Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
Db 928 AGATTCACATTTCAATTGACACGAGAGAGATTTTACGTGACTCAGCCCTTGGACCGA 987
Qy 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
Db 988 GAAAGAAAGATGCAATATGTTTTTATGCAAGTTGCAAGATGATGAGTAAACCACTT 1047
Qy 321 SerTyrProLeuGluIleHiseValLysValLysAspIleAsnAspAsnProProthCys 340
Db 1048 TCAATATCCGTGGAATATCTGTAAGATTAAAGATTATATATATCACTCAATCTGT 1107
Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluLysLeuGluLysAsnSerIleGly 360
Db 1108 CCGTACCACTAACCGATTTTGAAGTCCAGAGATATACGATGGGTAAACAGTATGGG 1167
Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrIleAsnSerPheLeuAsnTyrArg 380
Db 1168 ACCCTACTGACATGACAGGATGAGAAATATCTCCAAACAGTTTTCTTAACTACAGG 1227
Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
Db 1228 ATGTGTGAGCAAACTCCCAAACTTCCATGATGAGATCTTCTTAAATCCAAACCTAAGCT 1287
Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
Db 1288 GGAATGTTACAGTTAGCTAAACAGTCTCTTGAAGAGACAGATCTCTCAGTACACTTA 1347
Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
Db 1348 ACCATAGAGGTGTCTGACAAAGATTTCAAGACCCCTTTGTTGTGCAAAATCAACTTAT 1407
Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
Db 1408 GATTTCAATATATGATGATCCCATCTTTGAAAAATCAGATTTATGAAACCTGACTTGTCT 1467
Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db 1468 GAAAGACAAACATTTGGGTCCACCATCTTAACCATCCAGGCCACTGATGCTGATAGCCA 1527
Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLysGlyAspSerGluGlyValGlu 500
Db 1528 TTTACTGGGAGTTCTAAATTTCTGTATCATATCATATAAGGAGACAGTGGAGCGCTG 1587
Qy 501 GlyValAspThrAspProIshThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
Db 1588 GGGGTTCACAGATCCCATCCCATCCACAGCGAGATGCTAAATTAAGACCTCTGAT 1647
Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
Db 1648 TTTGAAACACAGCTGTTTCCAAATTTGTCTTCAAGCAGAAATCTCTGAGCCTTAGTG 1707
Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
Db 1708 TTTGTGTGAGATGATCAATGCAATGTTCTTTGCAAGTTCCAGCTTATTTGTGACAGATGTG 1767
Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
Db 1768 AATGAGCACCTCAATTTTCCCAACAGTATTCAGAGGAAAGTCAAGTAGAGATGATGCT 1827
Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
Db 1828 ATAGGACATTAAGTGGGCAATGATGCTGCCAAGATCCAAAGTCTGACATAGCTAT 1887
Qy 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyLysIlePhe 620
Db 1888 TCACGTGAGGAGACCAAGAGGTTGGCTTAAATATGACACAGCTGATCTGTGATCTTT 1947
Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrTyrGluValAlaAlaThr 640
Db 1948 AGTGTGCTTCATTTGACAGAGAGCGGAGATCCATATCGGGTACAAAGTGTGGCCACA 2007
Qy 641 GluValGlyLysSerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspVal 660
Db 2008 GAAGTGGGGGGCTTCTTAACTCTGTCTCAGATCTCACTGATCTTATAGATGTG 2067
Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db 2068 AATGACAAACCTCCAGGCTAGCCAGAACTACACGGGCTTGTCTTCTGCCATCCCTC 2127
Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheAspGly 700
Db 2128 AGTGCACCTGGAGATCTCATTTTCAGGCTACAGATGATATATAGCACTTAATTCGGGGT 2187
Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
Db 2188 CCCCATTTTACATTTTCCCTCGGCACTGAGAACTTACAAACACATCGGAAATTTCCAA 2247
Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluValArgAlaTyr 740
Db 2248 ATCAATATGATCATATCCGACATGTCTTACAGGACACACAGACTTTTGGAGAGGGGTAT 2307
Qy 741 ValValLeuIleArgIleAsnAspGlyLysArgProProLeuGlnGluGlyIleValSerLeu 760
Db 2308 GTGCTTATGATTCGATCAATGATGAGGGGTGGCCACCTTGGAGGACATTTTCTTTA 2367
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db 2368 CCAAGTTACATCTCGCATGTTGTGTGAGAGAGATTTTCCGGCAGCAGGTTCACAGACT 2427
Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db 2428 GGGATACCCCATGTGGGATGAGTGGATGATCTGTAACCAACCTTCTGTGATTTGTT 2487

Oy	801	llelleleuuAvalValPheIleKqrlleYLYSAspLYSGlyYSAspPenVlgu	820
Dd	2468	ATAATTTTAGCGAGTGTGTTTTCCGCATAAGAAGATTAAGCGAAATATTCTGA	254
Oy	821	SerAlaglnAlasSerGlualylsProLeuAtgSer	832
Dd	2548	AGTGCTCAAGCATCTGAAGTCAAACCTCTGAAGAAGC	2593

RESULT 4
US-09-833-263-1076

; Sequence 1076, Application US/09833263
; Patent No. US20020110547A1

APPLICANT: Wang, Aijun

APPLICANT: Stolck, John A.

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND

FILE REFERENCE: 210121.471C12
CURRENT APPLICATION NUMBER: IIS/09/833.263

```

; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093

```

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 3.0
;
; SEO ID NO 1076

```

```

; LENGTH: 3345
; TYPE: DNA

```

US-09-833-263-1076

Alignment Scores:

Pred. No.:	0	length:	3345
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-025-380-1081 (1-832) X US-09-833-263-1076 (1-3345)

QY	1	MetIleLeuGlnAlaHisLeuHisSerLeuQySleuLeuMetLeuTyrlleuAlaMrpGly	20
Db	88	ATGAATACTTCAGGCCCATCTTCACTCCCGTGTCTTATGCTTATATTGGCACTCGA	147
QY	21	TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrglu	40
Db	148	TATGGCCAAAGAGGGAGAGTTTAGTGAACCCCTGAACCCATGACATTTCTATTATGA	207
QY	41	GlyGlnLysProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe	60
Db	208	GGCCCAAGAACCCAGTCAATTTATATTCGAGTTTAAAGCCAAATCTCTCGTGGACTTTT	267
QY	61	GluLeuThrGlyGlyLuthrAspAsnIlePheValIleGlyAlaGlyGlyLeuLeuTyrryr	80
Db	268	GACTACTCTGGGGAGACAGACAAACATTTTGTGATGAACGGAGGAGACTCTGTATTAC	327
QY	81	AsnArgAlaLeuAspArgLuthrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp	100
Db	328	AACAGAGCCTTGGACAGGAAACAAAGATCTACTCAAACTCTCAGGTTGGACGCTCGAC	387
QY	101	AlaAsnGlyIleIleValGlyGlyProValProIleThrIleGlyValLysAspIleAsn	120
Db	388	GCTAATGGAAATTAAGTGAAGGCTCCAGTCCCTATACCACTAAAGTGAAGGACATCAC	447
QY	121	AspAsnArgProThrPheLeuGlnSerLysTyrgluGlySerValArgIleAsnSerArg	140
Db	448	GACATTCGACCCACGTTTCTCCAGCTCAAAAGTAGAAGGCTCAGTTAAGGACGACTCTGC	507
QY	141	ProGlyLysProPheLeuTyrrValAsnAlaThrAspLeuAspAspProAlaThrProAsn	160
Db	508	CCAGGAACCCCTCTGTATGTCAATGCCACAGACTTGATATCCGGCCACTCCCAAT	567

QY	161	GLYGLNLEUPLYRTRGIIILEVALILEGILEUPEPROMETILEASNSENVALLMETYRPH	180
DB	568	GGCCAGCTTATATACAGATTGTCATCCACTTCCATGATCAACATGCACTACTT	627
QY	181	GLINILEASNSENVSTHRGIALALESERLEUTHRAGSLGYSERGINILEUASN	200
DB	628	CAGATCAACAAACAAAGGAGGCATCTCTTACCAGAGGAGATCTCAGAAATGAAT	687
QY	201	PROALALYSANPROSETRYRANLEUVALILESERVALLYSAPMETGYGLINSE	220
DB	688	CTCTGCTAAGATCTCTTCTTAATCTGGTATCTCAGTGAAGACATGGGAGGCCAAGT	747
QY	221	GLIANSSEPHESERAPTHRTTRSERVALASPILEILEVALIHRGIANENILETPLYS	240
DB	748	GAGAAATTCCTCAGTGATACCACTCTGTGGATATCATAGTACAGAAATATTGGAAA	807
QY	241	ALAPROLYSEPROVALGIUMETVALIGLIANSSETRHNASPPOHISPROILEYSLIETHR	260
DB	808	GCACCAAAACCTGTGGAGATGGTGGAAAATCAACGATCTCACCACCCATCAAAATCACT	867
QY	261	GLINVALIARGTRPASNAPPROGLIYALGLINTYRSEULEUVALNPLYSGULYSEU	280
DB	868	CAGGTGGGGTGAATGATCCCGGGTCACAATATCTCTAAGTACAAAGGAAGCTGCCA	927
QY	281	ARGHERPROHESERILEASPGINGULYASPILEYRYVALIHRGINPROLEUASPAR	300
DB	928	AGATTTCCCATTTCAATYACCGAAGAGATATTTACGTGCTCAGCCCTTGGACCA	987
QY	301	GLIUGLIYASAPALATRYVALIPIHETRYVALIVALIYASAPGIUTRYGLYSPROLEU	320
DB	988	GAAAGAAAAGATGCAATATGTTTTTAAGCATTCGACATTCGGAAGATGATCGGAAAAACCTT	1047
QY	321	SERYTRPROLEUGIUILEHISVALIYSEVALLYSAPILLEASNSAPANPROPTHCY	340
DB	1048	TCAATATCCGCTGGAATTCATGTAAAGTTAAAGATTTATGATTAATCCACCTACATGT	1107
QY	341	PROSEPROVALIHRVALIHRGIIUVALIHRGIIUASNGIURPRLEUGIYANSERILEGLY	360
DB	1108	CCGTCACACGATACCGTATTTGAGGTCCAGAGATATACGACTGGGTAAACGATATCGG	1167
QY	361	THIRLEUTHRALHISAPARGASPGIUGLIUASNTHRALAANSERPHLEUASNTRYARG	380
DB	1168	ACCTTACTGCAKATGACAGGATGAAGAAAATATCTCCACAGTTTCTTAACTACAGG	1227
QY	381	IILEVALIUGINTHRPROLYSEUPROMETASPGIYLEUPELEUILEGINTHTRYALA	400
DB	1228	ATTGTGGAGCAACTCCCAAACTTCCCATGATGATCTTCTTCAATCCAAACCTATGCT	1287
QY	401	GLYMETLEUGIULEUVALIYEGINSETRLEUVALYSEGINASPHTRPROGINTRYRANLEU	420
DB	1288	GGAAATGTTACCGTTAGCTTAAACACTCTTGAAGAAGCAAAATACTCTCTCAGTACACTTA	1347
QY	421	THIRILEGIUVALISERAPLYSAPPHELYSTRLEUCYSPHEVALIGINILEASNVALILE	440
DB	1348	ACGATTAAGGTGTGTGCACAAAGATTTCAAGACCTTTGTTTGTGCAAAATCAAGTTATT	1407
QY	441	ASPILEASNSAPGINILEPROILEPHEGLIULYSESERAPTRYGIYANLEUTHRLEUALA	460
DB	1408	GATATCAATGATATCGATCCCATCTTTGAAAAATCAGATTATGAAACCTGACTCTTGCT	1467
QY	461	GLIUSAPHRANENILEGYSERTHIILEUETHRIILEGINIATHRAPALASPGIUPRO	480
DB	1468	GAAGACACAAACATTTGGGTCCACCATCTTAAACCATCCAGGCCACTGATGCGATGAACCA	1527
QY	481	PHETHRGISESERLYSILEUETHYRIILEILEYSGIYASPSERGIUGIYARGLEU	500
DB	1528	TTTATCGGAATTTTAAATATCTGTATCATATCTTAAGGAGACAGTGAAGGACGCTGT	1587
QY	501	GLIYALASPTHRASPPROHISTHRASNTHGLIYRYVALIILEILEYSLYSPROLEUASP	520
DB	1588	GGGGTTGACACAGATCCCATACCAACACCGGATATGTCATTAATTAAGCCCTTGAT	1647
QY	521	PHEGIUTHRALIYALISERASNILEVALIPIHETRYVALIAGLIUASNPROGLIUPROLEUVALI	540

Db 1648 TTGAAACAGACGAGTGTTCACACATGTTCTTCAAGAGAGAAATCCGACCTCTAGTG 1707
 Qy PheGlyVallylsrYrAsnAlaSerSerPheAlaIysPheThrLeuIleValThrAspVal 560
 Db 1708 TTGTGTGAGAGTACATGACATGATTTCTTTCCAAAGTTCCAGCTTATTTGACAGAGTG 1767
 Qy AsnGluAlaProGlnPheSerGlnIleValPheGlnAlaIysValSerGluAspValAla 580
 Db 1768 AATGAGACCTCAATTTTCCCAACAGTATTCAGAGAAAGTCAGTGAAGAGTGTAGCT 1827
 Qy 581 ILeGlyThrIysValGlyAsnValThrAlaIysAspProGluGlyLeuAspIleSerTyr 600
 Db 1828 ATAGGCACTAAAGGAGGCAATGTGACTGCCAAGATCCAGAAAGCTCGACATAGGCTAT 1887
 Qy SerLeuArgGlyAspThrArgGlyTyrLeuIysIleAspHisValThrGlyGluIlePhe 620
 Db 1888 TCACCTAGGGAGACACAGAGGTTGGCTTAAATTAACAACGAGACTGTGAGATCTTT 1947
 Qy SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 Db 1948 AGTGTGCTCATTTGACAGAGAACCGGAAAGTCATTCGGGTACAGAGTGTGGCCACA 2007
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAGTAGGGGGGTCTTCTTAAGCTGTGTCAAGATTCACATGCTTATGAGATGTG 2067
 Qy 661 AsnAspAsnProProArgLeuAlaIysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACAAACCTCCAGAGCTTAGCCAAAGACACAGGAGCTTGTCTTCTGCAATCCCTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnIleLeuPheArgGly 700
 Db 2128 AGTGCACCTGGAAGTCTCATTTTGAGGCTACTGATGATGACGACTTATTTCCGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerIys 720
 Db 2188 CCCCATTTTCAATTTTCCCTCGCAGCAGTGAAGCTTACAAACGACTGGAAAGTTTCCAA 2247
 Qy 721 ILeAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaThr 740
 Db 2248 ATCAATGATGACTACAGCCGACCTGTCTACAGGACACAGACTTGAAGAGGCGCTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db 2308 GTGTCTTGAATCCGACATGATGAGGGGTCCGCCACCTTGAAGGACTTTTCTTTA 2367
 Qy 761 ProValThrPheCysSerCysValGlyGlySerCysAspPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAGTTTACATTTGCAAGTGTGTGAGAGAAAGTTGTTCCGGCAGCAGAGTCAACAGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGAATATCCCATGCGGAGCATGCGAGTGTGATATGTGTACCAACCTTCTGAGTATTTGT 2487
 Qy 801 ILeIleLeuAlaValAlaPheIleArgIleIleIysAspIysGlyIysAspAsnValGlu 820
 Db 2488 ATATATTTTACAGTGTGTATTTATCCGATTAAGAGATTAAGGCAAGATATATTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValIysProLeuArgSer 832
 Db 2548 AGTGCTCAAGCACTGAAGTCAAACTGTGAAGAC 2583
 RESULT 5
 US-09-880-107-3288
 ; Sequence 3288, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3288
 ; LENGTH: 3345
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U07969
 US-09-880-107-3288
 Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-025-380-1081 (1-832) x US-09-880-107-3288 (1-3345)
 Qy 1 MetIleLeuGlnAlaHisIleLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 88 ATATATCTTCAAGCCCATCTTCACTCCCTGTCTTCTTATATTTGAGCAACTGGA 147
 Qy 21 TyrGlyGlnGluGlyIysPheSerGlyProLeuIysProMetThrPheSerIleTyrGlu 40
 Db 148 TATGCCAAGAGGGGAAGTTTATGAGACCTTGAAACCCATGACATTTTCTATTTATGAA 207
 Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheIysAlaAsnProProAlaValThrPhe 60
 Db 208 GGCACAAAGACCGAGTCAATATATTCAGTTTAAGGCCAATCTCTGCTGACCTTT 267
 Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuTyrTyr 80
 Db 268 GAACATACTGGGAGACAGACCAATATTTGTATGATGAACGGAGGAGACTTCTGTATAC 327
 Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 328 AACAGACCTTGAACAGGGAACAAGATCTACTACATCTCCAGGTGCGACCTGAGAC 387
 Qy 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValIysAspIleAsn 120
 Db 388 GCTAATGGAATTTATGAGAGGCTCCAGTCCCTATCACCATAGAGAGGACATCAAC 447
 Qy 121 AspAsnArgProThrPheLeuGlnSerIysTyrGlyGlySerValArgGlnAsnSerArg 140
 Db 448 GACAAATCGACCCAGTTCCTCCAGTCAAGATGAAGAGGCTCAGTAAGGCGAACTCTCGC 507
 Qy 141 ProGlyIysProPheLeuTyrValAlaAsnAlaThrAspLeuAspProAlaThrProAsn 160
 Db 508 CAGGAAGAGCCCTTCTTGTATGATCAATGCAACAGACTGTGATATCCGGCACTCCCAAT 567
 Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 568 GGCACCTTATATCAAGATGTCATCCAGCTTCCCATATCAACATGTCATGATCTTT 627
 Qy 181 GlnIleAsnAsnIysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 628 CAGATCAACAACAACCGGAGCAGCATCTCTTACCCGAGAGGAGATCTCAGGAATGGAAT 687
 Qy 201 ProAlaIysAsnProSerTyrAsnLeuValIleSerValIysAspMetGlyGlyGlnSer 220
 Db 688 CTGCTAAGATCTCTTCTTATATCTGAGATCTGATGAGACATGGAAGGCGCAGAGT 747
 Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrPlys 240

Db 748 GAGAAATTCCTTCAGTGAATACCAATCTGTGATATCATAGTACAGAGAAATATTTGGAAA 807
 Qy 241 AAlaProlyserProvalGluMetValGluAsnSerThrAspProHisProIleValIleThr 260
 Db 808 GCACCAAAACCTGTGGAGATGTGTGAAAACCACTGATCTCTCAACCCATCAAAATACT 867
 Qy 261 GlnValArgThrAsnAspProGlyValGlnIleThrSerLeuValAspIleValGluLeuPro 280
 Db 868 CAGGTGGGGTGAATGATCCCGGTGCACAAATTTCTTGAATGACAAAGAGACCTGCCA 927
 Qy 281 ArgPheProPheSerIleAspGlnGluValAspIleThrValThrGlnProLeuAspArg 300
 Db 928 AGATTCCCATTTTCATTTGACCAAGAGATATTTACGAGCTCAGCCCTTGAGCCGA 987
 Qy 301 GlnGluIlysaAspAlaThrValPheThrAlaValAlaIlysaAspGluThrGlyIlysaProLeu 320
 Db 988 GAAAGAAAAGATGATATGTTTTTATGCAAGTTCAAAGATGATGAGTGGGAAAACCACTT 1047
 Qy 321 SerThrProLeuGluIleHisValIlysaValIlysaAspIleAsnAspAsnProThrCys 340
 Db 1048 TCATATCCGCTGGAAATTCATGTAAAGTTAAAGATTAATGATATCCCATCACTCATCT 1107
 Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1108 CCGTACACAGTAAACCGATTTGAGGTCCAGAGATGAAACGATGGGTAAACAGATATCGGG 1167
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnThrArg 380
 Db 1168 ACCCTTACTGCATGACACAGGATGAAAGAAATACTGCAACATTTTCTAACTACAGG 1227
 Qy 381 IleValGlnGlnThrProIleLeuProMetAspGlyLeuPheLeuIleGlnThrThrAla 400
 Db 1228 ATTGTGAGCAAACTCCCAACTTCCTCCATGATGATGATCTCTTCAATCCAACTATGCT 1287
 Qy 401 GlyMetLeuGlnLeuAlaIlysaGlnSerLeuIlysaGlnAspThrProGlnThrAsnLeu 420
 Db 1288 GGAATGTATACAGTATGCTAAACAGTCTCTGAAAGACAAAGATACCTCCATCAACACTTA 1347
 Qy 421 ThrIleGluValSerAspIlysaAspPheIleThrLeuCysPheValGlnIleAsnValIle 440
 Db 1348 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCTTTGTTTGTGCAATCAAGTATTT 1407
 Qy 441 AspIleAsnAspGlnIleProIlePheGluIlysaSerAspIlysaIleLeuThrLeuAla 460
 Db 1408 GATATCAATGATTCAGATCCCATCTTTGAAAATCAAGTTATGAAAACCTGACTTGGCT 1467
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1468 GAAGACAAACATTTGGGTCCACCATCTTAACATCCAGGCCACTGATGATGAGCA 1527
 Qy 481 PheThrGlySerSerIlysaIleLeuThrHisIleIleIlysaAspSerGluIlysaGlu 500
 Db 1528 TTTACTGGAGATTTCTAAATCTGATCATATCAATAAAGGAGACAGAGAGGCGCTG 1587
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyThrValIleIleIlysaProLeuAsp 520
 Db 1588 GGGGTGACACAGATCCCATACCAACCCGATATGATATTAATAAAACCTCTTGAT 1647
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheIlysaGluAsnProGluProLeuVal 540
 Db 1648 TTTGAAACAGACAGCTGTTTCAACATTTGTGTTCAAGAGAAATCCCTGAGCCCTAGTG 1707
 Qy 541 PheGlyValIlysaThrAsnAlaSerSerPheAlaIlysaPheThrLeuIleValThrAspVal 560
 Db 1708 TTTGTGTGAAGTCAATGACAAAGTCTTTTCCCAAGTTCACAGCTTATTTGTACAGAGTG 1767
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIlysaValSerGluAspValAla 580
 Db 1768 AATGAGCACTCATTTTCCCAACACGTATTTCAAGGAAAGTCAAGAGATGATGACT 1827
 Qy 581 IleGlyThrIlysaValGlyAsnValThrAlaIlysaAspProGluGlyLeuAspIleSerThr 600
 Db 1828 ATAGGCACTTAAGTGGCAATGTGATCTGCCAAGGATCCAGAAAGTCTGTGACATATGAT 1887

Qy 601 SerLeuArgGlyAspThrArgGlyThrPheLeuValAspHisValThrGlyGluIlePhe 620
 Db 1888 TCATGAGGGGAGACACAAAGGTTGGCTTAAATGATCACACTGATGATGATCTTT 1947
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProThrArgValGlnValAlaIleThr 640
 Db 1948 AGTGTGGCTCATTTGACAGAGAAAGCCGGAAGTCCATATGCGGTATCAAGGTGGCCACA 2007
 Qy 641 GlnValGlyIlysaSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2008 GAAATGAGGGGCTTCTTCAAGCTCTGTCTCAGAGTTCACCTGATCTTATGATGATG 2067
 Qy 661 AsnAspAsnProProArgLeuAlaIlysaAspIlysaThrGlyLeuPhePheCysHisProLeu 680
 Db 2068 AATGACACACCTCCAGGCTAGCCAGACTACACGGGCTGTCTTCTGCACTCCCTC 2127
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2128 AGTGCACCTGGAAGTCTCATTTTTCGAGGCTACTGATGATGATCAGCATTAATTTCCGGGT 2187
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAspAspThrProGluValSerIlysa 720
 Db 2188 CCCCATTTTACATTTTCCCTCCGACGTGAAAGCTTACAAACGACTGGGAAATTTCCAA 2247
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaIlysa 740
 Db 2248 ATCAATGTACTCTATGCTCCGACACTGTCTACCGGACACACAGACTTTAGAGAGAGGCTAT 2307
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db 2308 GTCGTCTTGAATCCGATCATATGATGGGGGTGGCCACCTCTGAAAGGCAATTTGTTCTTA 2367
 Qy 761 ProValThrPheCysSerCysValGluIlysaSerCysPheArgProAlaGlyHisGlnThr 780
 Db 2368 CCAATTAATCTGCAATGCTGTGTGTGAAAGATTTGTTCCGCGACAGAGTCAACAGACT 2427
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2428 GGGATACCCACTGTGGGCAATGCGAGTGTGATCTGTGACACACCTTCTGATGATTCGT 2487
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleIlysaAspIlysaGlyIlysaAsnValGlu 820
 Db 2488 ATAAATTTACAGTGTGTTATCCCATTAAGAAAGATTAAGCAAAAGATTAATGTTGAA 2547
 Qy 821 SerAlaGlnAlaSerGluValIlysaProLeuArgSer 832
 Db 2548 AGTGTCTCAAGCATCTGAAATCAAAACCTCTGAGAGAC 2583

RESULT 6
 US-09-962-832-118
 Sequence 118, Application US/09962832
 Patent No. US20020110821A1
 GENERAL INFORMATION:
 APPLICANT: Ebner, Reinhard
 TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signat
 TITLE OR INVENTION: Sets
 FILE REFERENCE: 689290-74
 CURRENT APPLICATION NUMBER: US/09/962,832
 CURRENT FILING DATE: 2001-09-25
 PRIOR APPLICATION NUMBER: US/60/235,077
 PRIOR FILING DATE: 2000-09-25
 PRIOR APPLICATION NUMBER: US/60/235,280
 PRIOR FILING DATE: 2000-09-25
 NUMBER OF SEQ ID NOS: 259
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 118
 LENGTH: 3654
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-962-832-118
 Alignment Scores:

Pred. No.: 0 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 2
 Query Match: 99.58% Indels: 0
 DB: 10 Gaps: 0

US-10-025-380-1081 (1-832) x US-09-962-832-118 (1-3654)

QY 1 MetIleuGlnAlaHisLeuHisSerLeuCyseuMetLeuTyrLeuAlaThrGly 20
 Db ATGATCTTCAAGCCCATCTTCACTCCCTGTCCTCTTATGCTTTATTTGGCACTGGA 152
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuYsPrometThrPheSerIleTyrGlu 40
 Db TATGGCCAGAGGGGAAAGTTTAGTGAACCCCTGAACCCAGATTTTCTATTATGA 212
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLysAlaAsnProProAlaValThrPhe 60
 Db GGCCAGAACCCGAGTCAATTAATTAATTCAGTTTAAGCCCAATCCCTCGTGTACTTTT 272
 QY 61 GluLeuThrGlyGlyIleThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80
 Db GAATTAATCTGGGAGAGACAGACATATTGTGATGAAGACGGAGGACTTCTGTATTAC 332
 QY 81 AsnArgAlaLeuAspArgGlyIleThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db AACGAGCCCTTGGACAGGAAACAAGATCTACTACATCTCCAGTTTGCAGCCCTGGAC 392
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db GCTATATGAATTAATAGTGAAGGCTCCAGTCCCTTCAACATTAAGAGGAGCATCAAC 452
 QY 121 AspAsnArgProThrPheLeuGlnSerIleTyrGlyIleSerValArgGlnAsnSerArg 140
 Db GACATATGACCCAGCAGTCTTCTCACTCAAGTCAAGAGCTCAAGTAAAGCAGAACTCTGC 512
 QY 141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db CAGAGAAAGCCCTTCTGTATGTCATGCAATGCCACACCTGATGATCCGGCCACTCCCAT 572
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db GGCAGAGCTTATATACAGATTTGTCATCCAGCTCCCATGATCAACATGTCATGTACTT 632
 QY 181 GlnIleAsnLeuLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db CAGATCAACAACAACGAGGAGCATCTCTTAACGAGAGGATCTCGAGAAATTAAT 692
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 Db CTGCTAAGAAATCTTCTTAATCTGTGATCTCAATGAAAGACATGGAGGCCAAGT 752
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
 Db GAGAATTCCTTCAAGTATCCACATCTGTGATATCATAGACAGAGAAATATTGGAAA 812
 QY 241 AlaProLysProAlaGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db GACACCAAAACCTTGAGAGATGAGTGAACATCAATCTCAACCCCATCAAAATACT 872
 QY 813 GACACCAAAACCTTGAGAGATGAGTGAACATCAATCTCAACCCCATCAAAATACT 872
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db CAGGTGGGTGGAGATGATCCCGGTCACATATTCCTTAAGTGAACAAGAGAGCTGCA 932
 QY 873 CAGGTGGGTGGAGATGATCCCGGTCACATATTCCTTAAGTGAACAAGAGAGCTGCA 932
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
 Db AGATTCCCATTTCAATTGACAGAGAGATATTAAAGTACTGAGCCCTTGAGACGA 992
 QY 933 AGATTCCCATTTCAATTGACAGAGAGATATTAAAGTACTGAGCCCTTGAGACGA 992
 QY 301 GlnGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 Db GAAGAAAAGATGATATGTTTATGAGATGGCAAGAAGATGATGAGGAAAACCACTT 1052
 QY 993 GAAGAAAAGATGATATGTTTATGAGATGGCAAGAAGATGATGAGGAAAACCACTT 1052

QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 Db TCATATCCGCTGGAAATTCATGTAAGATTAAAGATATTATGATATTCACCTACATGT 1112
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db CCGTCACACAGTAACCGATATTGAGTCCAGAGATCAACGACTGGGTAAACATATCGGG 1172
 QY 361 ThrIleuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db ACCCTTCTGCACTGACAGAGATGAAGAAATATCTGCCAACGTTTCTTAATCTACAGG 1232
 QY 1173 ACCCTTCTGCACTGACAGAGATGAAGAAATATCTGCCAACGTTTCTTAATCTACAGG 1232
 QY 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db ATTGTGAGCAAACTCCCAAACTTCCATGATGATCTTCTTAATCCAACTATAGCT 1292
 QY 1233 ATTGTGAGCAAACTCCCAAACTTCCATGATGATCTTCTTAATCCAACTATAGCT 1292
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
 Db GGAATGTTACAGTTAGTAAACAGTCCCTGAAGAGCAAGATATCTCTCAGTACAACTTA 1352
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysePheValGlnIleAsnValIle 440
 Db ACGATAGAGGTGCTGCAAGAGATTCACAGCCCTTGTGTGTGCAAAATCAACGTATT 1412
 QY 1353 ACGATAGAGGTGCTGCAAGAGATTCACAGCCCTTGTGTGTGCAAAATCAACGTATT 1412
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db GATATCAATGATCAGACCCCATCTTGAATAATCAGATTATGGAACCTGACTTGCT 1472
 QY 1413 GATATCAATGATCAGACCCCATCTTGAATAATCAGATTATGGAACCTGACTTGCT 1472
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db GAAAGACAAACATTTGGTCCACCATCTTAACATCCAGGCCCACTGATGAGAACCA 1532
 QY 1473 GAAAGACAAACATTTGGTCCACCATCTTAACATCCAGGCCCACTGATGAGAACCA 1532
 QY 481 PheThrGlySerSerIleLeuTyrHisIleIleLysGlyAspSerGluArgLeu 500
 Db TTATCTGGAGTCTTAATAATCTGTATCATATATTAAGAGAGACATGAGGAGCGCTG 1592
 QY 1533 TTATCTGGAGTCTTAATAATCTGTATCATATATTAAGAGAGACATGAGGAGCGCTG 1592
 QY 501 GlyValAspThrAspProHisIleThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db GGGGTGACACAGATCCCATCCATCCACACCGGATATGCTAATTAATAAAGCCCTTGAT 1652
 QY 1593 GGGGTGACACAGATCCCATCCATCCACACCGGATATGCTAATTAATAAAGCCCTTGAT 1652
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db TTGAAACACAGACCTGTTTCCACATGTGTTCAAGACAAATAATCTGAGCCCTGATG 1712
 QY 1653 TTGAAACACAGACCTGTTTCCACATGTGTTCAAGACAAATAATCTGAGCCCTGATG 1712
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db TTTGGTGTGAAGTACAAATGCAAGTCTTTTGGCAAGTTCACGCTTATTTGACAGAGTG 1772
 QY 1713 TTTGGTGTGAAGTACAAATGCAAGTCTTTTGGCAAGTTCACGCTTATTTGACAGAGTG 1772
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db AATGAAGCACTCAATTTTCCCAACAGTATTCACACGAAAGTCACTGAGGATGTAGCT 1832
 QY 1773 AATGAAGCACTCAATTTTCCCAACAGTATTCACACGAAAGTCACTGAGGATGTAGCT 1832
 QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db ATAGGCACTAAAGTGGCAATGATGACTGCCAAGATCCAGAAAGTCTGACATAGCTAT 1892
 QY 1833 ATAGGCACTAAAGTGGCAATGATGACTGCCAAGATCCAGAAAGTCTGACATAGCTAT 1892
 QY 601 SerLeuAspGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
 Db TCACTGAGGGGAGACACAAGAGGTGTGCTTAATAATTGACACATGCTGTGATGATCTTT 1952
 QY 1893 TCACTGAGGGGAGACACAAGAGGTGTGCTTAATAATTGACACATGCTGTGATGATCTTT 1952
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValValAlaThr 640
 Db AGTGTGGCTCCATTGACAGAGAGAGCCGAGAGTCCATATCGGTTACAGTGTGGCCACA 2012
 QY 1953 AGTGTGGCTCCATTGACAGAGAGAGCCGAGAGTCCATATCGGTTACAGTGTGGCCACA 2012
 QY 641 GlnValGlyLysSerSerLeuSerSerValSerGlnPheHisLeuIleLeuMetAspVal 660
 Db GAAGTAGGGGGGCTTCTTGAAGCTCTGTGATGAGTCACTGATTCCTTATGATGATG 2072
 QY 2013 GAAGTAGGGGGGCTTCTTGAAGCTCTGTGATGAGTCACTGATTCCTTATGATGATG 2072
 QY 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db AATGACAAACCTTCCAGGCTGAGCAAGACTACACGGGCTGTGTTCTTCCCATCCCTC 2132
 QY 2073 AATGACAAACCTTCCAGGCTGAGCAAGACTACACGGGCTGTGTTCTTCCCATCCCTC 2132
 QY 681 SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly 700

Db 2133 AGTGCACCTGGAAGTCTCATTTTCGAGGCTACTATGATGATGACCTTATTTCCGGGT 2192
Qy 701 ProhibitinPhosphoserineGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
Db 2193 CCCATTTCATTTTCCCTCGCGAGGAGGCTTACAAAACGACTGGGAAGTTTCCAAA 2252
Qy 721 ILAASNGLYTHRHSAIAAGLEUSERTHRAGHISHPHEGILUGIUAARGLAATYR 740
Db 2253 ATCAATGAGTACTGACCCGACGTCTACACGACACAGAGTTTGAGAGAGGAGATAT 2312
Qy 741 ValValLeuLeuLeuArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
Db 2313 GTCCTCTTGATCCGCATCANTANGAGGGGGTGGCACTCTTGGAAAGCATTTGTTCTTTA 2372
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
Db 2373 CCAAGTTCATCTTCGACATTTGTGTGGAAGAAAGTTGTTCCGGCAGCAGGTCACCAACT 2432
Qy 781 GlyLeuProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
Db 2433 GGGATACCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2492
Qy 801 ILelleuAlaValIlePheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
Db 2493 ATAAATTTAGCAGTGTGTTTATCCGCATAAAGAGATTAAGCAAAAGATTAATGTTGAA 2552
Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
Db 2553 AGTCTCAGCATCTGAAGTCAAACCTCTGAGAAGC 2588

RESULT 7

US-09-880-107-3807
Sequence 3807, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-NO
CURRENT APPLICATION NUMBER: US/09/880.107
PRIORITY FILING DATE: 2001-06-14
PRIORITY APPLICATION NUMBER: US 60/211,379
PRIORITY FILING DATE: 2000-06-14
PRIORITY APPLICATION NUMBER: US 60/237,054
PRIORITY FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3807
LENGTH: 3654
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 X83228
US-09-880-107-3807

Alignment Scores:

Pred. No.: 0
Score: 4303.00
Percent Similarity: 99.76%
Best Local Similarity: 99.52%
Query Match: 99.58%
DB: 10

Length: 3654
Matches: 828
Conservative: 2
Mismatch: 0
Indels: 0
Gaps: 0

US-10-025-380-1081 (1-832) x US-09-880-107-3807 (1-3654)

Qy 1 MetIleuGlnAlaHisLeuHisSerLeuGlySerLeuMetLeuTyrLeuAlaThrGly 20
Db 93 ATGATACTTCAGGCCCATCTTCCTGCTCTTCTTATGCTTTATTTGGCACTGGA 152
Qy 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40

Db 153 TATGGCAAGAGGGGAAGTTAGTGAACCCCTGAACCATGACATTTCTATTATGAA 212
Qy 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPhe 60
Db 213 GGCAGAAAGACCGAGTAAATTAATTCAGGTTTAAAGCCCAATCTCTGCTGACTTT 272
Qy 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluLysGluGlyLeuLeuTyrTyr 80
Db 273 GAACTTAAGTGGAGACAGACAACTATTGTTGATGAAGGAGGAGGAGGAGGAGGAGGAG 332
Qy 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
Db 333 AACAGAGCCTTGGACAGGAGAAACAGATCTACCAATCTCCAGGTTGAGCCCTGGAC 392
Qy 101 ALAASNGLYIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
Db 393 GCTAAATGGAATTAATGAGAGGAGTCCAGTCCCTATACCATTAAGAGTGAAGAGCATCAC 452
Qy 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlySerValArgIleAsnSerArg 140
Db 453 GACATCGACCCCATCTTCTCCAGTCAGAGTGAAGAGGCTCAGTAAAGGAGAACTCCGC 512
Qy 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
Db 513 CCAAGAAAGCCCTTCTTGATGATGCAATGCCACAGACTGGATGATCCGGCCTCCCAAT 572
Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db 573 GGCACCTTAATTAACAGATGTCATCCAGTCCCATGATGATCAACATGATGATCTT 632
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
Db 633 CAGATCAACAAACAAACGGAGCATCTCTTACCAGAGGAGATCTCAGGAATGAAAT 692
Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
Db 693 CTGCTAAGATCTCTTCTTAATCTGATGATCTCAGTCAAGTGAAGAGGAGGAGGAGGAG 752
Qy 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
Db 753 GAAATATCTCTGATGATACACATCTGTGATATCATGATGATGATGATGATGATGATGAT 812
Qy 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
Db 813 GCACCAAACTGTGAGATGATGAGTGAAGTCAACGATCTCAACCCCAACAAATCACT 872
Qy 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
Db 873 CAGGTGGGTGGAATGATCCCGGTGCACAATATTCCTAGTTGACAAAGAGAGCTGCCA 932
Qy 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
Db 933 AGATTCCTCATTTCAATGATGACGAGAGGATTAATTAATGATGATGATGATGATGATGAT 992
Qy 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
Db 993 GAAAGAAAGATGATGATGATTTTATGAGTTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1052
Qy 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProProThrCys 340
Db 1053 TCATATCCGTGGAATATCTGTAAGTAAAGATTAATTAATGATGATGATGATGATGATGAT 1112
Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
Db 1113 CCGTCAACAGTAAACGATTTAGGTCACAGAAATGAAGATGAGTGAAGTGAAGTGAAGTGAAG 1172
Qy 361 ThrLeuThrAlaHisAspArgGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
Db 1173 ACCCTTAATCAATGACAGGAGTGAAGAAATTAATGATGATGATGATGATGATGATGATGAT 1232
Qy 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400

Db 1233 ATTGTGACGAACTCCCAAACTTCCATGATGACTCTTCTTAATCCAACTATGCT 1292
 Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnIleuValLeu 420
 Db 1293 GGATGTAGAGTAGTAAACACTCCTTGAAGAGCAAGATCTCCTCTGATCAACTTA 1352
 Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuLysPheValGlnIleAsnValIle 440
 Db 1353 ACCGATACAGGTGTGACAAAGATTGACAGCCCTTGTGTTTGGCAATCAAGTTATT 1412
 Qy 441 AsnIleAsnAspGlnIleProIlePheGlnLysSerAspThrGlyAsnLeuThrLeuAla 460
 Db 1413 GATATCATGATTCAGACCCCATCTTGAATAACATTAATGAAACCTGACCTTGCT 1472
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1473 GAAGACACAAACNTGGGTCCACCATCTTAACATCCAGGCCATGATGCTGATGAACCA 1532
 Qy 481 PheThrGlySerSerLysIleLeuThrIleIleLysGlyAspSerGluGlyArgLeu 500
 Db 1533 TTTTACTGGAGATTCTAAATTTCTGTATCATATCAATAAGGAGACAGTGAAGGACGCTG 1592
 Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1593 GGGGTGACACAGATCCCATACCAACACCGGATATGCTAATTAATAAGCCCTTGAT 1652
 Qy 521 PheGluThrAlaIleValSerAsnIleValPheLysAlaGluAspProGluProLeuVal 540
 Db 1653 TTTGAAACACAGCTGTTTCCACATTTGTGTTCAAGACAAAGAACTTGAGCCTTGATG 1712
 Qy 541 PheGlyValLysThrAsnAlaSerSerPheAlaLysPheThrLeuIleValIleAspVal 560
 Db 1713 TTTGGTGTGAGTACATGCAAGTTCTTTGGCAAGTTCAAGCTTATGTGACAGATGCTG 1772
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1773 AATGAAGCACCCTCAATTTCCCAACAGTATTCACACCGAAAGTCAGTGAAGTGTGATG 1832
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db 1833 ATAGGCACTAAAGTGGGCAATGACTGACGCAAGATCCAAAGGTCCTGACATAGACTAT 1892
 Qy 601 SerLeuArgLysAspThrArgLysTyrPheLysIleAspHisValThrGlyLysIlePhe 620
 Db 1893 TCAGTAGGGAGACACAAAGGTTGGCTTAATTAATGACCACGTAAGTGTGATGCTTT 1952
 Qy 621 SerValAlaProLeuAspArgLysAlaGlySerProTyrArgValGlnValIleAlaThr 640
 Db 1953 AGTGTGGCTCCATTTGACAGAAAGCCGAAAGTCCATATCGGGTACAAAGTGTGGCCACA 2012
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 2013 GAAGTAGGGGGGCTTCTTGAAGCTCTGTGTCAAGTTCCACTTGATCTTAAGAGTGTG 2072
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 2073 AATGACAAACCTCCCAAGGCTAGCAAGGACTACACGGGCTGTGTTCTTCTGCAATCCCTC 2132
 Qy 681 SerIleProGlySerLeuIlePheGluAlaThrAspAspAspGlnIleLeuPheArgGly 700
 Db 2133 AGTGACACCTGGAGTCTCATTTTCCAGGCTACGATGATGATATCAGCACTTAATTTCCGGGT 2192
 Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspThrProGluValSerLys 720
 Db 2193 CCCCATTTTACATTTTCCCTCGGAGGAGAGCTTCAAAAACAGCTGGGAAAGTTTCCAAA 2252
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaLys 740
 Db 2253 ATCAATGATCATCATGCGGACGCTGTCTACAGGACACAGAGTTTGAGAGAGAGGATAT 2312
 Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 Db 2313 GTGCTCTGATTCGATCAATGATGAGGGGTGCGCACCTTGAAGGCAATGTTTCTTTA 2372

Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2373 CAGTTATCAATCTTCGCACTTGTGTGAAAGAGATTGTTCCGGCAGCAGGTTCAACACT 2432
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2433 GGGATACCCCACTGTGGGCGATGCGAGTGTGATCTGCTGACCACTTCTGTGATTTGT 2492
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2493 ATATTTTAGCAGTGTGTTTATCCCATTAAGAGCATTAAGCAAAAGATAATGTTGAA 2552
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2553 AGTCTCAACACTGTAAGTCAAACCTCTGAGAAC 2588.

RESULT 8

US-10-025-380-1086
 / Sequence 1086, Application US/10025380
 / Publication No. US20020182191A1
 / GENERAL INFORMATION:
 / APPLICANT: Xu, Jiangchun
 / APPLICANT: Lodes, Michael J.
 / APPLICANT: Secrist, Heather
 / APPLICANT: Benson, Darin R.
 / APPLICANT: Meagher, Madeleine Joy
 / APPLICANT: Stolk, John A.
 / APPLICANT: Wang, Tongtong
 / APPLICANT: Jiang, Yuguang
 / APPLICANT: Smith, Carole L.
 / APPLICANT: King, Gordon E.
 / APPLICANT: Wang, Aijun
 / APPLICANT: Clapper, Jonathan D.
 / APPLICANT: Skelky, Yashir A. W.
 / APPLICANT: Fanger, Gary R.
 / APPLICANT: Vedvick Thomas S.
 / APPLICANT: Carter, Darrick
 / TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 / TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
 / FILE REFERENCE: 210121.471C14
 / CURRENT APPLICATION NUMBER: US/10/025,380
 / NUMBER OF SEQ ID NOS: 1129
 / SOFTWARE: PasteSeq for Windows Version 4.0
 / SEQ ID NO 1086
 / LENGTH: 2877
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 US-10-025-380-1086

Alignment Scores:

Pred. No.: 0 Length: 2877
 Score: 4208.00 Matches: 813
 Percent Similarity: 99.51% Conservative: 1
 Best Local Similarity: 99.39% Mismatches: 2
 Query Match: 97.38% Indels: 2
 DB: 9 Gaps: 1

US-10-025-380-1081 (1-832) x US-10-025-380-1086 (1-2877)

Qy 17 LeuAlaThrGly-----TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMet 34
 Db 421 CTGGCTAGCGGANTCCGGCCCGCACAAAGAGGAGATTGAGACCCCTGAAACCCATG 480
 Qy 35 ThrPheSerIleTyrGlyGluGlnGluProSerGlnIleIlePheGlnPheLysAlaAsn 54
 Db 481 ACATTTTCAATTTATGAAAGCCCAAGAACCGAATTAATATTCAGTTTAAAGCCAT 540
 Qy 55 ProProAlaValAlaThrPheGluLeuThrGlyGluThrAspAsnIlePheValIleGluArg 74
 Db 541 CTTCTGCTGTGACTTTTGAACCTAACTGGGAGAGACAGACAAATATTTGTGATAGAACGG 600


```

Db      2761 ACCCTTGGTGTGATTGTGATTAATTTAGACAGTTGTGTTATTCGCCATTAAGAGATTA 2820
Qy      815 G1YLSAspAnVal1G1uSer1aG1n1aSer1uVal1ySPProLeuAArgSer 832
      2821 GGCAAAATATATATGTGAAAGTGCTCAAGCATCTGAAGTCAAACTCTGAGAGC 2874

RESULT 9
US-09-922-217-1086
; Sequence 1086, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secret, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongcong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922.217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ. ID NOS: 1124
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1086
; LENGTH: 2877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-217-1086

Alignment Scores:
Pred. No.: 0 Length: 2877
Score: 4208.00 Matches: 813
Percent Similarity: 99.51% Conservative: 1
Best Local Similarity: 99.39% Mismatches: 2
Query Match: 97.38% Indels: 2
DB: Gaps: 1

US-10-025-380-1081 (1-832) x US-09-922-217-1086 (1-2877)
Qy      17 LeuAlaThrGly-----TyrGlyGlnGlnGlyLysPheSerGlyProLeuLysProMet 34
      421 CTGGCTAGCGGATCCGCGCGCCGACAAAGAGGGGAAGTTAGTGACCCCTGAAACCCTATG 480
Qy      35 ThrPheSer1LeuTyrGlnGlyGlnGluProSerGln1Le1PheGlnPheLysAlaAsn 54
      481 ACATTTTCTATTAATGAAGGCAAGAAACGAGTCAATTAATTCAGTTAAAGGCAAT 540
Qy      55 ProPro1aValThrPheGlnLeuThrGlyGlnGluThrAspAsn1LePheVal11GluArg 74
      541 CTCCTGCTGCTGACTTTTGAACCTACCTGCGGAGACACAAACATTTTGTGATGAACGG 600
Qy      75 GlnGlyLeuLeuTyrTyrAsnArgAlaLeuAspArgGlnThrArgSerThrAsnLeu 94
      601 GAAGGAGCTTCTGATATACAAACAGAGCCTTGACAGGAAACAAAGATCTACTCAATCTC 660
Qy      95 GlnValAlaAlaLeuAspAlaAsnGlyTLe1ValGlnGlyProValPro1LeuThr1Le 114
      661 CAGGTTGACCCCTCGACGCTAATGATTAATGATGAGGCTCAAGTCCCTATCACATA 720
Qy      115 GlnValLysAsp1LeaAspAsnArgProThrPheLeuGlnSer1YrGlnGlySer 134
      721 GAAGTGAAGACATCAACGACATCGACCACTTTTCTCAGTCAAAAGTGAAGGCTCA 780
Qy      135 ValArgGlnAsnSerArgProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAsp 154
      781 GTAAAGGCAAACTCTGCCACAGAAAGCCCTTCTTGATGTCAATGCCACAGACTGGAT 840

```

```

Qy      155 AspPro1aThrProAsnGlnGlnLeuTyrTyrGln1LeVal11GlnLeuProMet1Le 174
      841 GATCCGGCCACTCCCAATGCGCAGCTTTATTAACAGATTGTCAAGCTTCCCATATAC 900
Qy      175 AsnAsnValMetTyrPheGln1LeaAsnLysThrGlyAla1LeSerLeuThrArgGln 194
      901 AACAAATGTATGTATCTTTCAGATCAACCAAAACGGAGCCACTCTCTTACCCGAGAG 960
Qy      195 GlySerGlnLysLeuAsnProAlaLysAsnProSerTyrAsnLeuVal11LeSerValLys 214
      961 GGATCTCAGGAATGGAATCTGTCTAAGAACTCTCTATATCTGTGATCTCAGTGAAG 1020
Qy      215 AspMetGlyGlnGlnSerGlnAsnSerPheSerAspThrThrSerValAsp1LeVal 234
      1021 GACATGGAGGCGCAGATGGAATTCCTTACGATACCACTGCTGGAATATCATATGTC 1080
Qy      235 ThrGlnAsn1LeTyrLysAlaProLysProValGlnMetValGlnAsnSerThrAspPro 254
      1081 ACAGAGAAATATTTGAAAGCACCAAAACCTGTGAGATGTGTGAAACCTCACTGATCCT 1140
Qy      255 HisPro1LeLys1LeThrGlnValArgTyrAsnAspProGlyValGlnTyrSerLeuVal 274
      1141 CACCCCATCAAAATCACTCAGGCGGTGGAATGATCCCGTGCACAAATATCTTATGTT 1200
Qy      275 AspLysGlnLysLeuProArgPheProPheSer1LeAspGlnGlnGlyAsp1LeTyrVal 294
      1201 GACAAAGAAAGCTGCGCAAGATTCCTCAATTTTCAATTTGACAGAAAGCAATATTTACGTG 1260
Qy      295 ThrGlnProLeuAspArgGlnGlnLysAspAlaTyrValPheTyrAlaValAlaLysAsp 314
      1261 ACTCAGCCCTTGGACGAGAAAGAAAGATGATATGTTTTTATGCAAGTTGCAAGAGAT 1320
Qy      315 GlnTyrGlnLysProLeuSerTyrProLeuGln1LeHisValLysValLysAsp1LeAsn 334
      1321 GAGTACGGAAGAAACACTTTCATATCCGCTGGAATTCATGTAAAGTAAAGATATTTAAT 1380
Qy      335 AspAsnProProThrCysProSerProValThrValPheGlnValGlnGlnAsnGluArg 354
      1381 GATTAATCACTTCAATGCTCGTCAACAGTAACTGTTAGGTCCAGGAAATGAACGA 1440
Qy      355 LeuGlyAsnSer1LeGlyThrLeuThrAlaHisAspArgAspGlnGlnAsnThrAlaAsn 374
      1441 CTGGGTAAACGATTCGGGACCTTACTGCACATGACAGGATGAAGAAATATCTGCCAAC 1500
Qy      375 SerPheLeuAsnTyrArg1LeValGlnGlnThrProLysLeuProMetAspGlyLeuPhe 394
      1501 AGTTTCTTAAACTTACAGATGTGTGAGCAAACTCCCAAACTTCCCATGATGACTCTTTC 1560
Qy      395 Leu1LeGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysLysGlnAsp 414
      1561 CTATTCAAACCTATCTGAAATGTTTACAGTTAGCTTAAACAGTCTTGAAGAGCAAGAT 1620
Qy      415 ThrProGlnTyrAsnLeuThr1LeGlnValSerAspLysAspPheLysThrLeuCysPhe 434
      1621 ACTCTCAGTACAACTTAAACGATGAGGTCTCGAAGAAATTTCAAGACCTTTGTTTT 1680
Qy      435 ValGln1LeAsnVal1LeAsp1LeaAsnAspGln1LePro1LePheGlnLysSerAspTyr 454
      1681 GTCCAAATCAACGTTATGTTATCAATGATCAAGATCCCACTTTGAAAAATCAGATTAAT 1740
Qy      455 GlnAsnLeuThrLeuAlaGlnAspThrAsn1LeGlySerThr1LeLeuThr1LeGlnAla 474
      1741 GGAACCTGACTCTTCTGTAAGACAAACATTTGGGTCCACCAATCTTAACCTCAAGGCC 1800
Qy      475 ThrAspAlaAspGlnProPheThrGlySerSerLys1LeLeuTyrHis1Le1LeLysGly 494
      1801 ACTGATGCTGATGAGCAATTTACTGGAGTTCTAAATCTGTATCATATCAATAAGGGA 1860
Qy      495 AspSerGlnGlyArgLeuGlyValAspThrAspProHisThrAsnThrGlyTyrVal11Le 514
      1861 GACAGTGAAGGAGCGCTGGGGGTGTGACACAGATCCCATACCAACACCGGATATGTGCATA 1920

```

QY 515 ILELYSVPROLEUASPHEGLUTHRALALAVASERANILEVALPHELYSALAGLU 534
 DB 1921 ATRRAAAAGCCTTGATTTTGAACAGACGCTTTTCCAACTGGTTCCTCAAGCGAA 1380
 QY 535 ASNPRLUPLUPEUVALPHEGLYVALLYSTYRANALASERSERPHEALALVSPHETHR 554
 DB 1981 AATCTAGGCTCTAGTGTGGGTGAGTACAAATCAATCTCTTTTGGCAAGTTCACG 2040
 QY 555 LEUULEVALTHRASPVALAENGUVALAProGlnPheSerGlnHISVALPHEGLNALALYS 574
 DB 2041 CTTATTGTGACAGATGGAATGAAGACCTCAATTTTCCAAACGATTCCTCAAGCGAA 2100
 QY 575 VALSERGLUASPVALALALIEGLYTHRIFYVALGLYASNAVALTHRALALYASPPROGLU 594
 DB 2101 GTCACTGAGATGATGCTATGAGCACTTAAGTGGGCAATGTGACTGCCCAAGATCCGAA 2160
 QY 595 GLYLEUASPILLESERYSERLEUARGLYASPPHARGLYTPLEULYVILEASPHIS 614
 DB 2161 GGCTGGAACATAAGCTATTCTAGAGGGAGACACAAGAGGTGGCTTAAATTCAGCAC 2220
 QY 615 VALTHRGLYGLUTIERPHESEVALALAPROLEUASPPROGLUAGLYSERPROTYRARG 634
 DB 2221 GTGACTGTGAGATCTTAAAGTGGCTCCATTTGACAGAGAGCCGGAAGTCCATATCGG 2280
 QY 635 VALGLNVALVALATHRGUVALGLYGLYSESERLEUASPPHARGLYTPLEULYVILEASPHIS 654
 DB 2281 GTACAAAGTGTGGCCACAGAAAGTGGGGGCTTCTTAACTGTGTGACAGATTCAC 2340
 QY 655 LEUULEUWETASPVALASNAASPPROBROHARGLEUALALYASPPYRTHGLYLEU 674
 DB 2341 CTGATCTCTTATGATGATGATGACACCTCCAGCTAGCAGCAAGATCCACGGGCTTG 2400
 QY 675 PHEPHECYHISAPROLEUASERVALAProGlySERLEUILEPHEGLUALATHRASPASP 694
 DB 2401 TTTCTTCCCATCCCTCAGACCTGGAAGTCTCATTTTCCAGGCTACTGATGATGAT 2460
 QY 695 GLNHSLEUPHEARGLYPROHISPHETHRPHESERLEUGLYSERGLYSELEUGLNAS 714
 DB 2461 CAGACCTTATTTCCGGGCTCCCATTTTTCATTTTCCCTCGGCATGGAAAGCTTCAAAA 2520
 QY 715 ASPTRPGLUVALSERLYVILEASNGLYTHRHISALARGLEUSERTHRARGHISTHRASP 734
 DB 2521 GACTGGGAGATTTCCAAATCAATGATCATGCTCCGACTGTACAGGCAACAGAC 2580
 QY 735 PHEGLUVALARGALATYRVALVALLEUILEARGILEASNAEPGLYVARGPROLEU 754
 DB 2581 TTTGAGAGAGAGGGGTATGCTGTGATCCGATCATGATGAGGGGCTGCCACCTTG 2640
 QY 755 GLUGLYILEVALSERLEUPROVALTHRPHESERCYSEVALGLYSECYSPHEARG 774
 DB 2641 GAAGGCAATGTTTCTTTACCGATTACATTTCTGCACTGTGTGGAAGAAAGTTTCCG 2700
 QY 775 PROHLAGLYHISGINTHGLYILEPROTHRVAGLYMETALVALGLYILELEUENH 794
 DB 2701 CCACAGGATCCACAGACTGGATACCCACTGTGGGCAATGCAATGGATCTCTGACC 2760
 QY 795 THRILEUVALILEGLYILEILEUVALVALPHEILEARGILEVLYASPLYS 814
 DB 2761 ACCCTTCTGGTGTATGATATTTTACAGCTGTGTATCCCATTAAGAAAGATPAA 2820
 QY 815 GLYLYSAPNAENVALGLUSERALAGLNASERGLUVALLYSPROLEUARGSER 832
 DB 2821 GGCAAAAGTATATGTTGAAGTGTCTCAAGCATCTAAAGTCAAACTCTGAGAAAGC 2874

RESULT 10

US-09-833-263-1086
 ; Sequence 1086, Application US/09833263
 ; Patent No. US20020110547A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1086
 ; LENGTH: 2877
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(2877)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-263-1086

Alignment Scores:
 Pred. No.: 0
 Score: 4208.00
 Percent Similarity: 99.51%
 Best Local Similarity: 99.39%
 Query Match: 97.38%
 DB: 10
 Matches: 2877
 Mismatches: 813
 Conservative: 1
 Indels: 2
 Gaps: 1

US-10-025-380-1081 (1-832) x US-09-833-263-1086 (1-2877)

QY 17 LEUVALATHRGLY-----TYRGLYGLNGLUGLYLYSPHESEGLYPROLEUYSPPROMET 34
 DB 421 CTGGCTGAGCGAGCCCGCGCCGACAAAGAGGGAAATTTAGTGAGCCCTCGAAACCAG 480
 QY 35 THRPHESERILETYRGLUGLYGLNGLUPROSEGLNILELEPHEGLNPELYSALASNA 54
 DB 481 ACATTTTCTATTATATGAGGCCAAGAACGAGCAATATATATCCAGTTTAAAGGCAAT 540
 QY 55 PROVALVALATHRPHGLUETHRGLYGLUTHRASPANILEPHEVALILEGLUARG 74
 DB 541 CTCTGCTGTGACTTTTGAATCACTAAGGGGAGACAGACAACTATTTGTGATGAAAGCG 600
 QY 75 GLUGLYILEUENHLYRTHRANARGALALEUASPPARGLUTHRPSESETHRHISANLEU 94
 DB 601 GAAGGACTTCTGATTAACAACAGAGCTTGGACAGGAAACAAAGATCTACTCAATCTC 660
 QY 95 GLNVALALALEUASPPALASNGLYILELEVALGLUGLYPROVALPROILETHRILE 114
 DB 661 CAGCTTGACCCCTTGACCGTATGATATGATGAGGGTCCATCTTACACCATTA 720
 QY 115 GLUVALLYSAPILLEASNAASPPROTHRPHLEUGLINSERYTYRGLUGLYSER 134
 DB 721 GAAGTGAAGACATCAACAGCAATCGAACCCCTTCTCCAGTCAAAAGTACGAAGGCTCA 780
 QY 135 VALARGLNASERLRPARGLYLYSPROPHLEUETHRVVALASNAATHRASPLEUASP 154
 DB 781 GTAAGGACGAACCTCTGCCCGAAGAGCCCTTTGTATGTCAATGCAACAGACTGGAT 840
 QY 155 ASPPROVALATHRPROASNGLYGLNLEUETHRYTYRGLNILEVALILEGLINLEUPROMET 174
 DB 841 GATCCGGCCACTCCCAATGGCCGCTTATATACAGATTTCTACAGCTTCCAGTATC 900
 QY 175 ASNAASNAWETRYRPHGLNILEASNAENLYTHRGLYVALILESERLEUTHRARGLU 194
 DB 901 AACAAATGTYATGATCTTTCAGATCAACAACAAACGAGGACCATCTCTTACCCGAGAG 960
 QY 195 GLYSEGLNGLUENHNPVALALYASNPSESETRHASNLEUVALILESERVALLYS 214
 DB 961 GATCTCAAGAAATGATCTCTGTAAGAACTCTCTATATATGTGGATCTCAAGTGAAG 1020
 QY 215 ASPMERCGLYGLYGLNGLUASNPSESETRHASPETHRTHSERVALAPILILEVAL 234
 DB 1021 GACATGGAGGCCAGACTGAGATTTCTTCAAGATACCAACATCTGGATATCATATGTC 1080
 QY 235 THRGLUASNILETRPYLSALAPROLYSPROVALGLUWETVALGLUASNPSESETRHASP 254

Db 1081 ACAGAGATATTTGGAAAGCAACCAAACTGTGAGATGTTGAGAAAACCTCACTGATCT 1140
 QY 255 HisProIleuYsIleThrGlnValArgTTrpAsnAspProGlyAlaGlnTyrSerLeuVal 274
 Db 1141 CAACCCATCAAAATCACTCAGTGCCTGGATGATATCCCGGTGACAAATATCTTTAGTT 1200
 QY 275 AspIlyGluYLeuProArgPheProPheSerIleAspGlnGluYAspIleTyrVal 294
 Db 1201 GACAAAGAGAAAGCTGCAGAGATTCCTCAATTCATGACCAAGAGAGATATTACGTG 1260
 QY 295 ThrGlnProLeuAspArgGluGluYLeuAspAlaTyrValPheTyrAlaValAlaYAsp 314
 Db 1261 ACTAGCCCTTGGACCGAGAAAGAAAGATATCTTTTATGAGATTGCAAGAT 1320
 QY 315 GluTyrGlyYsProLeuSerTyrProLeuGlnIleHisValYsValYsAspIleAsn 334
 Db 1321 GAGTACGGAAGAACCTTTCTATATCCCTGGAAATTCATGTAAAGTTAAAGATATTAT 1380
 QY 335 AspAsnProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArg 354
 Db 1381 GATATATCACTCATGTCCGTACAGTAACTGATATTGAGGTCCAGAGAAATGACGA 1440
 QY 355 LeuGluYsAsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsn 374
 Db 1441 CTGGGTAAAGATACGGGACCTTATCTGCACATGACAGGATGAAGAAAATACGCCAAC 1500
 QY 375 SerPheLeuAsnTyrArgIleValGluGlnThrProYsLeuProMetAspGlyLeuPhe 394
 Db 1501 AGTTTCTTAAACTACAGAGATTTGTGACCAATCCCAACTTCCCATGAGATGACTTTC 1560
 QY 395 LeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaYsGlnSerLeuYsYsGlnAsp 414
 Db 1561 CTAATCAAACTATGCTGAGATGTACAGTTAGCTAAACAGTCTTGAAGAACAGAT 1620
 QY 415 ThrProGlnTyrAsnLeuThrIleGluValSerAspYsAspPheYsThrLeuCysPhe 434
 Db 1621 ACTCTCTAGTAACTTAAAGATAGAGTGTCTGACAAAGATTTCAAGACCTTTGTTT 1680
 QY 435 ValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluYsSerAspTyr 454
 Db 1681 GTCCAAATCAACGTATTGATATCAATGATCAATCCCATCTTTGAAAATCAAGATTAT 1740
 QY 455 GlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIleGlnAla 474
 Db 1741 GGAACCTGACTCTGTGCTGAAGACACAACATTTGGGTCCACCATCTTAACATCCAGGCC 1800
 QY 475 ThrAspAlaAspGluProPheThrGlySerSerIlyIleLeuTyrHisIleIleYsGly 494
 Db 1801 ACTGATGCTGATAGCCATTTACTGGGAGTTCTTAAATTCGTATCATATCATTAAGGGA 1860
 QY 495 AspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThrGlyTyrValIle 514
 Db 1861 GACAGTGAAGGAGCCCTGGGGGTGTAACAGATCCCATACCAACACCGGATATGTATA 1920
 QY 515 IleYsYsPheProLeuAspPheGluThrAlaValSerAsnIleValPheYsAlaGlu 534
 Db 1921 ATTAAGAAAGCCTCTGATTTTGAAGACAGAGCTGTTTCCAAATGTGTTCAAGACGAA 1980
 QY 535 AspProGluProLeuValPheGlyValYsTyrAsnAlaSerSerPheAlaYsPheThr 554
 Db 1981 AATCCGAGGCTCTAGAGTTGGTGAAGTAACTCAAGTCTTTTGGCAAGTTCACG 2040
 QY 555 LeuIleValThrAspValaAsnGluAlaProGlnPheSerGlnHisValPheGlnAlaYs 574
 Db 2041 CTTATTTGTGCAATGTGAATGAGCACTCAATTTTCCCAACAGATATCCAAAGGCAAA 2100
 QY 575 ValSerGluAspValaIleIleGlyThrYsValGlyAsnValThrAlaYsAspProGlu 594
 Db 2101 GTCAGTGAAGGATGATGATAGGCACTAAAGTGGGCAATGTGACTGCCAAGATCCAGAA 2160
 QY 595 GlyLeuAspIleSerTyrSerLeuArgGlyAspThrArgGlyTyrLeuYsIleAspHis 614
 Db 2161 GGTCTGACATTAAGTATTCACTGAGGGGAGACAAAGAGTGTGGCTTAAATTTGACAC 2220

QY 615 ValThrGlyGluIlePheSerValAlaProLeuAspArgGluAlaGlySerProTyrArg 634
 Db 2221 GTACCTGTGAGATCTTTAGTGTGCTCATTTGACAGAGAACCGGAAGTCAATCCG 2280
 QY 635 ValGlnValAlaIleThrGluValGlyYsSerSerLeuSerSerValSerGluPheHis 654
 Db 2281 GTACAGTGGTGGCCACAGAAAGTGAAGGGGTCTTCTTAAAGCTGTGTGACAGTTCCAC 2340
 QY 655 LeuIleLeuMetAspValaAsnAspAsnProProArgLeuAlaYsAspTyrThrGlyLeu 674
 Db 2341 CTGATCTTATGATGTGAATGACAACTCCACAGGCTAGCCAAAGACTACAGGGCTTG 2400
 QY 675 PhePheCysHisProLeuSerAlaProGlySerLeuIlePheGlnAlaThrAspAsp 694
 Db 2401 TTCTTCTGCATCTCCCTCAGTGCACCTGGAAGCTCATTTTGAAGCTACTATGATGAT 2460
 QY 695 GlnHisLeuPheArgGlyProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsn 714
 Db 2461 CAGCACTTATTTGGGGTCCCATTTTACATTTTCCCTCGGAGTGAAGCTTACAAAC 2520
 QY 715 AspTyrGluValSerIlyIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAsp 734
 Db 2521 GACTGGAAAGTTCCAAATCAATGATGACTCATGCCGACTGTACAGGACACAGAC 2580
 QY 735 PheGluGluArgAlaTyrValValLeuIleArgIleAsnAspGlyGlyArgProProLeu 754
 Db 2581 TTGAGAGAGAGGCGATATGCTCTTGCATCCCATATGATGGGGTGGCCACCTTG 2640
 QY 755 GluGlyIleValSerLeuProValThrPheCysSerCysValGluGlySerCysPheArg 774
 Db 2641 GAAGCATTTGTTCTTTTACCATTTACTGCAAGTTGTGGAGAAAGATTTGTTCCGG 2700
 QY 775 ProAlaGlyHisGlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuThr 794
 Db 2701 CCAGCAGGTACACAGCTGGATACCCACTGTGGCATGGCATGGTGTATACGTGACC 2760
 QY 795 ThrLeuValIleGlyIleIleLeuAlaValPheIleArgIleYsYsAspYs 814
 Db 2761 ACCCTCTGCTGATTTGATTAATTTTGAAGTGTGTGTTATCCGATTAAGAGATTA 2820
 QY 815 GlyYsAspAsnValGluSerAlaGlnAlaSerGluValYsProLeuArgSer 832
 Db 2821 GGCAAAGATATATGTTGAAGTGTCAAGCATGTGAAGTCAACCTCTGAAGAC 2874

RESULT 11
 US-10-063-547-97
 ; Sequence 97, Application US/10063547
 ; Publication No. US20020182638A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerlitsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,547
 ; PRIORITY FILING DATE: 2002-05-02
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 97
 ; LENGTH: 2848
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 US-10-063-547-97

Alignment Scores:

Pred. No.: 3.24e-89 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 Gaps: 21

US-10-025-380-1081 (1-832) x US-10-063-547-97 (1-2848)

62 LeuThrglyglu-----ThraAspnlllePheValilegluarglu---Gly 76
 296 CTGTCAGGGGACTCAGGCAAGCACTGAGGGCCCATTTGCTATGAGATTCAGATTCTGGC 355
 77 LeuLeuTyrraAsnarglAlaLeuAspArglulThraArgSerThrHisAsnLeuGlnVal 96
 356 TTCTCTGCTGAGCAGGAGGCGCTTGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 415
 97 AlaAlaLeuAspAlaAsnGlyIleIleValgluIleProValProIleThrIleGluVal 116
 416 ACCCTGAGATGACAGATGACATGCTTGTGGGGTCCACAGCCCTGCTGTGCAGCTG 475
 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerIleTyrrgluIleValArg 136
 476 AAGGATGAGATGACAGAGTGCCTCATTTCTCAAGCCATCTACAGAGCTCGGCTAGC 535
 137 GlnAsnSerArgProglIyLysProPheLeuTyrrValAsnAlaThrAspLeuAspAspPro 156
 536 CGGGGTACAGAGGCTTGATGATCTTCTCTCTCTGAGGCTTCAAGCCGAGATGAGCA 595
 157 AlaThrProAsnGlnIleLeuTyrrgluIleValileGlnLeuProMetIleAsnAsn 176
 596 GGCAAGAGCAACTGGATCTTCATTCATCCAGCCAGAGCGGCTCAGCCAGCCCTTCC 655
 177 ValMetTyrrPheGlnIleAsnAsnIleThrGlyAlaIleSerLeuThrArgIleGlySer 196
 656 CCAGACATGTTCCAGCTGAGGCTCTGGGCTGGGCTGAGCCCTCAGCCCAAGAGAGAGC 715
 197 GlnIleLeuAsnProAlaLysAsnProSerTyrrAsnLeuValileSerValLysAspMet 216
 716 ACCAGCTTGAACAGCGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
 217 GlyIleGlnSerGlnAsnSerPheSerAspThrThrSerValAspIleIleValThrGln 236
 776 GGTACACAGGCCC--TCAGGCCACACAGGCACTGCACCGTGAAGATCTCCATCATAGAG 832
 237 AsnIleTyrrAspIleProLysProValGlnMetValGlnAsnSerThrAspProHisPro 256
 833 AGCACCTGGGTGCTGAGGCTTATCCACTGAGAGAGATCTCAAGTCTTATACCG 892
 257 IleLysIleThrGlnValArgIleAspAsnArgProglIyLysIleGlnTyrrSerLeuValAspLys 276
 893 CACCAACATGCGCCAGGTCACCTGAGAGGAGGAGTATGCACTTACACTG----- 943
 277 GlnLysLeuProArgPheProPheSerIleAspGlnGlnIleAspIleTyrrValThrGln 296
 944 GAGAGCAATCCCGCGGAGCCCTTGAAGTGAATGACAGAGGAAACCTTACGTGACAGA 1003
 297 ProLeuAspArgIleGlnLysAspAlaTyrrValPheTyrrAlaValAlaLysAspGlnTyrr 316
 1004 GAGCTGACAG 1063
 317 GlnLysProLeuSerTyrrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
 1064 GGGAG 1123
 1124 GTGCTATCTGCTCCCGCCGAGCCCAAGTCAAGCATCCCTGAGTCAAGTCAAGTCAAGTCA 1183
 337 ProProThrCysProSerProValThrValPheGlnValGlnIleGlnIleGlnIleGln 356
 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGlnIleGlnIleGlnIleGlnIle 376
 1184 ACTGAGTACTAGACTGTCTGACAGAGAGATGACAGATGCCCCCGGCTCCCAATTTCCAC 1243

377 LeuAsnTyrrArgIleValgluIleThrProLysLeuProMetAspIle---LeuPheLeu 395
 1244 GTTGTGTATCAGCTCTCGAGCCCTGAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1303
 396 IleGlnIleThrValIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 414
 1304 GTGAGCCCACTTCAGAGAGTGAAGCTGGGGGTCTCCACTCCGAGAGAGAGAGAGAGAG 1363
 415 ThrProGlnIleAsnLeuThrIleGlnIleValIleSerAspLysAsp-----PheLysThrLeu 432
 1364 ATCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnIleSer 452
 1424 TGTGAATGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1483
 453 AspTyrrGlnAsnLeuThrLeuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIle 472
 1484 CAGATTGGGCTTAAAGCTCTCCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1543
 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrrHis 490
 1544 ACAGCCATTCATGATCTGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1594
 491 IleIleLysGlyAspSerGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 510
 1595 ATTGAAG 1648
 511 GlyTyrrValIleIleLys-----LysProLeuAspPheGlnIleThrAlaAlaValSerAsn 528
 1649 GGGCATGTTAGACTGACAGCTGCAAGAACTCAAGTTAAGAGAGAGAGAGAGAGAGAGAG 1708
 529 IleValPheLysAlaGlnAsnProglIyLysProLysValPheGlyValLysTyrrAsnAlaSer 548
 1709 GT 1765
 549 SerPheAlaLysPheThrLeuIleValThrAspValAsnGlnIleProGlnPheSerGln 568
 1766 GGCACCGCCCAAGTGAATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1825
 569 HisValPheGlnAlaLysValSerGlnAspValAlaIleGlyThrIleValGlnVal 588
 1826 GAGAGTACAG 1885
 589 ThrAlaLysAspProglIyLysLeuAspIleSerTyrrSerLeuArgIleAspGly 608
 1886 CAGCCCTCCGAGCCCACTCAAG 1945
 609 TrpLeuLysIleAspHisValThrGlyIlePheSerValAlaProLeuAsp---Arg 627
 1946 TGGCTGTGATTAAGAAATCTCCGGGAGAGTCAACAGCCCAAGTCCCTGAGAGAGAGAG 2005
 628 GlnAlaGlySerProTyrrArgValGlnValAlaIleThrGlnValIleGlySerSerLeu 647
 2006 CAGCCTGGGAGACCTCAACAGGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2053
 648 SerSerValSerGlnPheHisIleLeuMetAspValAsnAspAsnProArgLeu 667
 2054 -----CTGACTCTTGCCCTGTGCTGCTCCCA----- 2080
 668 AlaLysAspTyrrThrGlyLeuPhePheCysHisAspProLeuSerAlaProGlySerLeuIle 687
 2081 -----TACCTCTGACAGCCCGGCAAGACCATGCTGATGAGTGTG 2119
 688 Phe---GlnAlaThrAspAspAspGlnHisLeuPheArgGlyProHisPheThrPheSer 706
 2120 AGTGAAG 2176
 707 LeuGly---SerGlySerLeuGlnAsnAspTrpGlnValSerLysIleAsnGlyThrHis 725
 2177 CTGTGTCCCAACCCAGGAG 2236
 726 AlaArgLeuSerThrArgHisThrAspPheGlnIleArgAlaTyrrValIleLeuIleArg 745


```

Db      2237 GCGTAACCTGACCTTGGCCCTGTCATGCGTGGAGCCAGCCATGAACATATCCCGCTGGTG 2296
Qy      746 ILeaNsApGlyIArgProPheLeuGluGlyIleValSerLeuProValThPheCys 765
Db      2297 GTCCAGCCCAATGCCCG-----ATGGGACAGCTCCGTGGTTCAGATATGCTGTGT 2347
Qy      766 SerCys--ValGluGlySerCysPheArgProAlaGlyHisGlnThGlyIleProThr 784
Db      2348 CGCTGCAACGTGGAGGGGCAATGCGATGCGCAAGGTGGCCCGCAGAGGCGCATGCCAGC 2407
Qy      785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleuAla 804
Db      2408 AAGCTGTGCGAGTGGGCGCATCTTGTAGGACCTGTGATGCAATAGGAACTCTTCATC 2467
Qy      805 ValValPheIleArgGlyIleValLeuAspGlySerIleValAspAsnValGluSerAlaGlnAla 824
Db      2468 CTCATTTTCACCCACTGGACCATGTCAAGAAAGAACCCGATCAACAGCAGACAGC 2527
Qy      825 SerGluValIys 828
Db      2528 GTGCCCTGAAG 2539

RESULT 12
US-10-063-616-97
; Sequence 97, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 97
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-616-97

Alignment Scores:
Pred. No.: 3,246-89 Length: 2848
Score: 908.50 Matches: 237
Percent Similarity: 50.51% Conservative: 159
Best Local Similarity: 30.23% Mismatches: 335
Query Match: 21.03% Indels: 53
DB: 9 Gaps: 21

US-10-025-380-1081 (1-832) x US-10-063-616-97 (1-2848)
Qy      62 LeuThrGlyGlu-----ThrAspAsnIlePheValIleGluArgGlu---Gly 76
Db      296 CTGTCAGGGAGCTGAGGCAAGGCACTGAGGGCCCATTTGCTATGATCCAGATTCCTGCG 355
Qy      77 LeuLeuTyTrpAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
Db      356 TTCCTGCTGTGACACGAGGGCCCTGGACCGAGAGGACGACAGTACCAAGCTACAGGTC 415
Qy      97 AlaAlaLeuAspAlaGlnGlyIleIleValGlnGlyProValProIleThrIleGlnVal 116
Db      416 ACCCTGAGAGATGCAAGGATGATGCTGTGGGGGTCAACAGCCCTGTGCTTGACAGTG 475
Qy      117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerIleTyArgIleSerValArg 136

```

```

Db      476 AAGATAGAGATGACCAAGTGCCCGCATTTCTCCAAAGCCATCTACAGAGCTCGGCTGAGC 535
Qy      137 GlnAsnSerArgProGlyIleProPheLeuTyValAsnAlaThrAspLeuAspPro 156
Db      536 CGGGGTACAGGCTGGCATCCCTTCCTTCCTTAGGCTTCAGACCGGATGAGCCA 595
Qy      157 AlaThrProGlnGlyGlnLeuTyTrpGlnIleValIleGlnLeuProMetIleAsnAsn 176
Db      596 GGCACAGCCAACTCGATCTTCGATTCACATCTTGAGCCAGAGGTCACAGCCAGCTTCC 655
Qy      177 ValMetTyRheGlnIleAsnAsnTyThrGlyAlaAlaIleSerLeuThrArgGlySer 196
Db      656 CAGACATGTTCCAGCTGAGGCGCTGGGCGGCTGGCCCTGACCCCAAGGGAGAGC 715
Qy      197 GlnGluLeuAsnProAlaValAsnProSerTyAsnLeuValIleSerValIleAspMet 216
Db      716 ACCAGCCTTACCAAGCCCTGAGAGGACCTACAGCTGTGGTACAGGTCAAGACATG 775
Qy      217 GlyGlyGlnSerGluAsnSerPheSerAspThrThSerValAspIleIleValThrGlu 236
Db      776 GTGACCAAGGCC--TCAGGCCACAGGCCCATGTCACGTTGAAAGTCTTCATATAG 832
Qy      237 AsnIleTrpLysAlaProLysProValGluMetValGlnAsnSerThrAspProHisPro 256
Db      833 AGCAGCTGGGTGTCCTTAGAGCCTATCCACTGCGAGAGATTCAAAGTCTATACCG 892
Qy      257 IleValIleThrGlnValArgTrpAsnAspProGlyValAlaGlnTySerLeuValAspLys 276
Db      893 CACCACATGGCCCAAGGTACCTGAGTGGGGGTGATGTGACATCACTG----- 943
Qy      277 GluTyLeuProArgPheProPheSerIleAspGlnGluGlyAspIleTyValThrGln 296
Db      944 GAAAGCATCCCGCGGAGCCCTTTGAAGTAAATGCAAGAGAAACCTTAGAGACAGA 1003
Qy      297 ProLeuAspArgGlnGluIleValAspAlaTyValPheTyAlaValAlaValAspGluTyR 316
Db      1004 GAGCTGACAGAGAGAGCCAGCGGTAGTACTCTGCTCAGGTGGGGGTGAGATTCAT 1063
Qy      317 GlyLysProLeuSerTyProLeuGluIleHisValIleValIleValIleAspAsn 336
Db      1064 GGGAGAGCATATGCGGCCCTCTGAGAGCTGACGTGCTGTGATGATGAGATGACAAAC 1123
Qy      337 ProProThrCysProSerProValThrValPheGluValGlnGlnAsnGluGly 356
Db      1124 GTGCTATCTGCTCCCTCCCGCTGACCCCAAGTACGACTGCTAGCTCAACAGGT 1183
Qy      357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGlnAsnThrAlaAsnSerPhe 376
Db      1184 ACTGAAGTACATGACGTGCTGACAGAGAGATGACAGATGCCCGGCTCCCATTTCCAC 1243
Qy      377 LeuAsnTyArgIleValGlnGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
Db      1244 GTTGTGATCATGCTCTGAGCCCTGAGCCCTGAGAGATGAGGGAGAGAGCTTCCAG 1303
Qy      396 IleGlnThrTyAlaGlyMetLeuGlnLeuAlaValGlnSerLeuIleLys--GlnAsp 414
Db      1304 GTGAGCCCACTTACAGAGTGTGACGCTGGGGGTCTCCACTCCGAGAGGCCAGAAC 1363
Qy      415 ThrProGlnTyAsnLeuThrIleGluValSerAspLysAsp-----PheTyThrLeu 432
Db      1364 ATCCGCTCTGCTGCTGCGCATGAGACCTGGCAGGGCGCAGAGGTGGCTTACAGACAGC 1423
Qy      433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
Db      1424 TGTGAAGTGAATGCCAGTACAGATATCAAGATCAACGCGCTGAGTTCATCAGCTTCC 1483
Qy      453 AspTyArgLysLeuLeuThrLeuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIle 472
Db      1484 CAGATTGGGCTATTAAGCTCTCCTGAGAGTGTGAGCCCGGAGCTGTGGGACATGCTA 1543
Qy      473 GlnAlaThrAspAlaAsp--GluProPheThrGlySerSerLysIleLeu--TyHis 490

```

[illegible]

```

RESULT 13
US-10-063-502-97
/ Sequence 97, Application US/10063502
/ Publication No. US20030023042A1
/ GENERAL INFORMATION:
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Geritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, Christopher J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3230R1C1
/ CURRENT APPLICATION NUMBER: US/10/063,502
/ CURRENT FILING DATE: 2002-05-01
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 170
/ SEQ ID NO 97
/ LENGTH: 2848
/ TYPE: DNA
/ ORGANISM: Homo Sapien
US-10-063-502-97

Alignment Scores:
Pred. No.: 3,246-89 Length: 2848
Score: 908.50 Matches: 237
Percent Similarity: 50.51% Conservative: 159
Best Local Similarity: 20.23% Mismatches: 335
Query Match: 21.03% Indels: 53
DB: 9 Gaps: 21

US-10-025-380-1081 (1-832) x US-10-063-502-97 (1-2848)

QY 62 LeuThrglyGlu-----ThrAspAsnIlePheValIleGluIuArgIu--Gly 76
DB 236 CTGTGAGGGGACTCATGCGCAGGCACTGAGGGCCATTGTCATGATGATCGAATTCCTGCG 355
QY 77 LeuLeuTYrTYrAsnArgIleLeuAspArgIuThrAspSerThrHisenLeuGlnVal 96
DB 356 TTCCTGCTGTGATCACCGAGGCCCTTGAGCCGAGAGGAGCAGCAGAGTACCAAGCTACAGGTC 415
QY 97 AlaAlaLeuAspAlaAsnGlyIleIleValIleGluIuProValProIleThrIleGluVal 116
DB 416 ACCCTGAGAGATCAGAGATGACATGCTGTGGGGGTCCACAGCTGTGCTTGACAGTG 475
QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerIleTYrGluIuGlySerValArg 136
DB 476 AAGATGAGGAATGACAGAGGTGCCCATTTCTCTCAAGCATCTACAGAGCTCGCTGAGC 535
QY 137 GlnAsnSerArgProGlyIuysProPheLeuTYrValAsnAlaThrAspLeuAspAspPro 156
DB 536 CGGGGTACACGAGCTGCGCATCCCTTCCTTCCTTGAGAGCTTCACACCGGATGAGCCA 595
QY 157 AlaThrProAsnGlyIleuLeuTYrGlnIleValIleGluLeuProMetIleAsnAsn 176
DB 596 GGCACAGCCCACTCGGATTTGCATTTCCATCTGACCGCAGGCTCCAGCCAGCCTTCC 655
QY 177 ValMetTYrPheGlnIleAsnAsnIysThrglyAlaIleSerLeuThrArgIuGlySer 196
DB 656 CCAGACATATTCACGTGAGGCTGGCGTGGGGGGCTGTGGCCCTGAGCCCAAGGGGAGC 715
QY 197 GlnGluLeuAsnProAlaLysAsnProSerTYrAsnLeuValIleSerValIysAspMet 216
DB 716 ACCACCTTGACACGAGCCCTTGAGAGAGACCTACAGCTGTGTACAGGTCAAGACATG 775
QY 217 GlyGlyGlnSerClnAsnSerPheSerAspTYrThrSerValAspIleIleValThrGlu 236
DB 776 GGTGACACGAGGCT--TCAAGCCACAGGCGCACTGCCACGTGGATGATCTTCATCATAGAG 832

```

QY 237 Asn1LeuTyrValAlaProLysProValGluMetValGluAsnSerThrAspProHisPro 256
 Db 833 AGCACTGGGTGCTCCCTAGAGCTTATCCACTGGCAGAGATCTCAAGTCTTATACCG 892
 QY 257 IleValIleThrGlnValArgThrAsnAspProGlyAlaGlnTyrSerLeuValAspLys 276
 Db 893 CACCAATGGCCAGTACCTGAGTGGGGTATGTGCACTTACCTG----- 943
 QY 277 GluLysLeuProAlaGlyPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGln 296
 Db 944 GAGAGCCATCCCCGGGACCTTTGAAAGTAAATGCAAGAGGAAACCTTACGTACAGCA 1003
 QY 297 ProLeuAspArgGluGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyr 316
 Db 1004 GAGCTGACAGAGAGCCAGCTGAGTACCTGTCCAGGCGGGCTCAGAAATTCCTCAT 1063
 QY 317 GluLysProLeuSerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
 Db 1064 GGGAGGAGACTATGGGGCCCTCTGAGCTGCACGTGTGTGATGATGAGAAATGACAC 1123
 QY 337 ProThrThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly 356
 Db 1124 GTGCTTATCTGCTCCCTCCCGAGCCCACTCAGCATTCCTGTGACTCAGTCCACAGGT 1183
 QY 357 AsnSerIleGluTyrIleuThrAlaHisAspArgAspGluAsnThrAlaAsnSerPhe 376
 Db 1184 ACTGAGTGAAGTACACTGTGCAGCAGAGATGCAAGATCCCGGCTCCCAATTTCCAC 1243
 QY 377 LeuAsnTyrArgIleValGluGlnThrProLysLeuProMetAspGly---LeuPheLeu 395
 Db 1244 GTTGTGTATGAGCTCCGAGCCCTGAGCCCTGAGCGATGGGTAAGGGGAGAGCTTCCAG 1303
 QY 396 IleGlnThrTyrAlaGluMetLeuGlnLeuAlaLysGlnSerLeuLys---GluAsp 414
 Db 1304 GTGAGCCCACTTACAGCAGTGTACCTGGGGGTGCTCCCACTCCAGCAGCCCAAC 1363
 QY 415 ThrProGlnTyrAsnLeuThrIleGluValSerAspLysAsp---PheLysThrLeu 432
 Db 1364 ATCTGCTTCTGGGTGGCCATGAGCTGAGCATGAGCCGAGAGGGGTGCTTACAGAGACG 1423
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
 Db 1424 TGTGAAGTCAAGCGCAGTCAAGATATGATATGATACGCCCTGTGATTCATCACTTCC 1483
 QY 453 AspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle 472
 Db 1484 CAGATTGGCCCTTAAAGCTTCCCTGAGAGATGTGAGCCCGGAGCTGTGTGGCCATGCTA 1543
 QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrHis 490
 Db 1544 ACAGCCATTGATGTGCTGACCTGAGCC-----GCCCTCCCGCTCAGATTTTGGC 1594
 QY 491 IleIleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
 Db 1595 ATTGAGGGGAGACACAGAGGAGCTTTGGCTGTGATGGGAGCA-----GACTCT 1648
 QY 511 GlyTyrValIleIleLys-----LysProLeuAspPheGluThrAlaAlaValSerAsn 528
 Db 1649 GGGAGTGTATGACTCAGACTCTGCAAAACCTCATATAGAGCATCCCAAGTCAATGAG 1708
 QY 529 IleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrAsnAlaSer 548
 Db 1709 GTGTGTGTGTGTGTGCGAGATGTGGCAAGCTGTG---GGCGCAGGCCCAAGCCCTCGA 1765
 QY 549 SerPheAlaLysPheThrLeuIleValIleThrAspValAsnGluAlaProGlnPheSerGln 568
 Db 1766 GCCACCGCCACGGGTGCTGT 1825
 QY 569 HisValPheGlnAlaLysValSerGlnAspValAlaIleGlyThrLysValGlyAsnVal 588
 Db 1826 GAGAGCTACAGAGCGCAGTGTCCCATCATGAGTCCCAAGCCGAGCTTCTTCTGTGACCATC 1885

QY 589 ThrAlaLysAspProGluGlnLysAspIleSerTyrSerLeuAspArgIleAspThrArgGly 608
 Db 1886 CAGCCCTCCGACCCCAATCAGCCGACCCCTCAGGTTCTCTTACATGATGATCAGAGGCC 1945
 QY 609 TrpLeuLysValIleAspHisValIleThrGlyGlnIlePheSerValAlaProLeuAsp---Arg 627
 Db 1946 TGGCTGTGCAATTAAGAAATTTCTCGGGAGATGACACCCGCCAGTCCCTGCAAGGGCC 2005
 QY 628 GluAlaGlySerProIleTyrArgValGlnValAlaIleThrGlnValGlyLysSerLeu 647
 Db 2006 CAGCTGGGAGACCTTACAGGCTCTTGTGAGGCCCAAGATACAGCC----- 2053
 QY 648 SerSerValSerGluPheHisIleLeuMetAspValAsnAspAsnProArgLeu 667
 Db 2054 -----CTGAGCTCTTGCCCTGTGCTCCCA----- 2080
 QY 668 AlaLysAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
 Db 2081 -----TACCTTGACACCCCGCCAGACCATGCTTGATCTGTG 2119
 QY 688 Phe---GluAlaThrAspAspArgGlnHisLeuPheArgGlyProHisPheThrPheSer 706
 Db 2120 AGTGAGCCAGCAGAGAACCCCGATCTGGCCAGTGGCGACGGTCC---TACAGCTTCCAC 2176
 QY 707 LeuGly---SerGlySerLeuGlnAsnAspTyrGluValSerLysIleAsnGlyThrHis 725
 Db 2177 CTGTGTCCCAACCCCAAGGTCGACAGGATGGCGGCTCAGACTCATGATGTTCCAT 2236
 QY 726 AlaArgLeuSerThrArgHisThrAspPheGluGlnArgAlaTyrValValLeuIleArg 745
 Db 2237 GCCTACCTCACTTGGCCCTGTCATTGGGTGAGCCAGTGAACATATCCCGGTGTG 2296
 QY 746 IleAsnAspGlyValArgProProLeuGlnGlyIleValSerLeuProValThrPheCys 765
 Db 2297 GTAGCCACAAATGCCAG-----ATGTGCAGCTCTGTGTGAGATGCTGTGT 2347
 QY 766 SerCys---ValGluLysSerCysPheArgProAlaGlyHisGlnThrGlyIleProThr 784
 Db 2348 CGGTGACAGTGGAGGGGAGCTGATCCGCAAGTGGCGCGCATGAAAGGCAGGCCAGC 2407
 QY 785 ValGlyMetAlaValAlaGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAla 804
 Db 2408 AAGCTTCGGCAGTGGGAGTCTTGTGAGCACCCTGTGAGCATAGAAATCTTCCATC 2467
 QY 805 ValValPheIleArgIleLysLysAspLysGlyLysAspAsnValGluSerAlaGlnAla 824
 Db 2468 CTCATTTCACCCATGCAATGTCAGAGAGAGAGACCCGATCAACCGACAGCAGC 2527
 QY 825 SerGluValLys 828
 Db 2528 GTGCCCTGAG 2539

RESULT 14
 US-10-227-884-187
 ; Sequence 187, Application US/10227884
 ; Publication No. US20030027988A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Gerritsen, Mary
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin I.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3530P1C79
 ; CURRENT APPLICATION NUMBER: US/10/227,884
 ; CURRENT FILING DATE: 2002-08-26

;; PRIOR APPLICATION NUMBER: 10/119,480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/065549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089905
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090691
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/095302
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095318
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095916
;; PRIOR FILING DATE: 1998-08-10
;; PRIOR APPLICATION NUMBER: 60/096146
;; PRIOR FILING DATE: 1998-08-11
;; PRIOR APPLICATION NUMBER: 60/096791
;; PRIOR FILING DATE: 1998-08-17
;; PRIOR APPLICATION NUMBER: 60/097986
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/098544
;; PRIOR FILING DATE: 1998-08-31
;; PRIOR APPLICATION NUMBER: 60/099596
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099598
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099803
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812

;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108801
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/11549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27

QY 549 SerPheAlaIysPheThrLeuIleValThrAspValAenGluAProGlnPheSerGln 568
Db 1766 GCCACCCGCGACGCTGCTAGTGTGAGAGATGATGCCACCCCGGACGTTGGACAG 1825
QY 569 HiValPheGlnAlaIysValSerGluAspValAlaIleGlyThrIysValGlyAsnVal 588
Db 1826 GAGAGCTACGAGGCGACATGTCGCCATGAGTCCCGCCAGCCGGCTCTTCCTGTCGACCATC 1885
QY 589 ThrAlaIysAspProGlnGlyLeuAspIleSerIleSerLeuArgIleAspThrArgIle 608
Db 1886 CAGCCCTCCACCCCTCAGCCGGAACCTCAGTTCCTCCAGTCAATGACTCAGAGGCG 1945
QY 609 ThrLeuIysIleAspHisValThrGlyIlePheSerValAlaProLeuAsp---Arg 627
Db 1946 TGGCTGCTGATGAAATTCCTCGGAGGTGACACCCGCGGCTCTGTCAGAGGGGCC 2005
QY 628 GluAlaIysSerProIleArgValGlnValAlaThrGluValGlyIleSerLeu 647
Db 2006 CAGCCTGGGACACCTTACAGGCTCTGTCGAGGCCGAGATACAGCC----- 2053
QY 648 SerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAspProArgLeu 667
Db 2054 -----CTGACTCTGCCCCCTGCTGCCCA----- 2080
QY 668 AlaIysAspIleThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
Db 2081 -----TACCTCTGCACACCCCGCCCAAGACCATGGCTTATGCTG 2119
QY 688 Phe---GluAlaThrAspAspAspGlnHisLeuPheArgIleProHisPheThrPheSer 706
Db 2120 AGTGGACCCGACGAGACCCGCGCTGCGCCAGTGGGACCGGTCC---TACAGCTTACCC 2176
QY 707 LeuGly---SerGlySerLeuGlnAspIleProGluValSerIleAsnGlyThrHis 725
Db 2177 CTGTGCTCCAACTCCCGCTGCAACGCGATTTGGCGCTCCAGACTCTCAATGCTTCCCAT 2236
QY 726 AlaArgLeuSerThrArgHisThrAspPheGluGlnArgAlaIleValIleuIleArg 745
Db 2237 GCTTACCTCACCCTTGGCCCTGATGGGTGAGGACCATGAACATTAATCCCGGTGTC 2286
QY 746 IleAsnAspGlyIleArgProProLeuGlnIleValSerLeuProValThrPheCys 765
Db 2297 GTCAGCCACCATGGCCAG-----ATGGGACGCTCGTGTTCAGTATGCTGTGT 2347
QY 766 SerCys---ValGlyIleSerCysPheArgProAlaGlyHisGlnThrGlyIleProThr 784
Db 2348 CGCTGCAACGTGGAGGCGCAGTGCATGCGCAAGTGGCCCGCATGAGGGCATGCCACG 2407
QY 785 ValGlyMetAlaValIleIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAla 804
Db 2408 AAGCTGCGGAGTGGGATCTTGTAGGCACTCTGTAGCAATGCAATGCTTCTCTCATC 2467
QY 805 ValValPheIleAspIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 824
Db 2468 CTCATTTTCAACCACTGACCATGTCAAGAAAGACCCGATCAACGACGACAGC 2527
QY 825 SerGluValIys 828
Db 2528 GTGCCCTGGAAG 2539

APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530P1C96
CURRENT APPLICATION NUMBER: US/10/230,163
CURRENT FILING DATE: 2002-08-28
PRIORITY APPLICATION NUMBER: 10/119,480
PRIORITY FILING DATE: 2002-04-09
PRIORITY APPLICATION NUMBER: 60/059113
PRIORITY FILING DATE: 1997-09-17
PRIORITY APPLICATION NUMBER: 60/062287
PRIORITY FILING DATE: 1997-10-17
PRIORITY APPLICATION NUMBER: 60/063549
PRIORITY FILING DATE: 1997-10-28
PRIORITY APPLICATION NUMBER: 60/064103
PRIORITY FILING DATE: 1997-10-31
PRIORITY APPLICATION NUMBER: 60/069873
PRIORITY FILING DATE: 1997-12-17
PRIORITY APPLICATION NUMBER: 60/078910
PRIORITY FILING DATE: 1998-03-20
PRIORITY APPLICATION NUMBER: 60/079294
PRIORITY FILING DATE: 1998-03-25
PRIORITY APPLICATION NUMBER: 60/079656
PRIORITY FILING DATE: 1998-03-26
PRIORITY APPLICATION NUMBER: 60/079728
PRIORITY FILING DATE: 1998-03-27
PRIORITY APPLICATION NUMBER: 60/081819
PRIORITY FILING DATE: 1998-04-15
PRIORITY APPLICATION NUMBER: 60/081955
PRIORITY FILING DATE: 1998-04-15
PRIORITY APPLICATION NUMBER: 60/082804
PRIORITY FILING DATE: 1998-04-22
PRIORITY APPLICATION NUMBER: 60/084441
PRIORITY FILING DATE: 1998-05-06
PRIORITY APPLICATION NUMBER: 60/085323
PRIORITY FILING DATE: 1998-05-13
PRIORITY APPLICATION NUMBER: 60/085579
PRIORITY FILING DATE: 1998-05-15
PRIORITY APPLICATION NUMBER: 60/086392
PRIORITY FILING DATE: 1998-05-22
PRIORITY APPLICATION NUMBER: 60/089532
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089538
PRIORITY FILING DATE: 1998-06-17
PRIORITY APPLICATION NUMBER: 60/089905
PRIORITY FILING DATE: 1998-06-18
PRIORITY APPLICATION NUMBER: 60/090472
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090557
PRIORITY FILING DATE: 1998-06-24
PRIORITY APPLICATION NUMBER: 60/090691
PRIORITY FILING DATE: 1998-06-25
PRIORITY APPLICATION NUMBER: 60/090695
PRIORITY FILING DATE: 1998-06-25
PRIORITY APPLICATION NUMBER: 60/091982
PRIORITY FILING DATE: 1998-07-07
PRIORITY APPLICATION NUMBER: 60/095302
PRIORITY FILING DATE: 1998-08-04
PRIORITY APPLICATION NUMBER: 60/095318
PRIORITY FILING DATE: 1998-08-04
PRIORITY APPLICATION NUMBER: 60/095916
PRIORITY FILING DATE: 1998-08-10
PRIORITY APPLICATION NUMBER: 60/096146
PRIORITY FILING DATE: 1998-08-11
PRIORITY APPLICATION NUMBER: 60/096791
PRIORITY FILING DATE: 1998-08-17
PRIORITY APPLICATION NUMBER: 60/097986
PRIORITY FILING DATE: 1998-08-26
PRIORITY APPLICATION NUMBER: 60/098544
PRIORITY FILING DATE: 1998-08-31
PRIORITY APPLICATION NUMBER: 60/099596


```

PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29

PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835

Alignment Scores:
Pred. No.: 3,24e-89 Length: 2848
Score: 908.50 Matches: 237
Percent Similarity: 50.51% Conservative: 159
Best Local Similarity: 30.23% Mismatches: 335
Query Match: 21.03% Indels: 53
DB: 9 Gaps: 21

US-10-025-380-1081 (1-832) x US-10-230-163-187 (1-2848)
QY 62 LeuThrglyGlu-----ThraSpaaniIlephValIIeqluArglu---gly 76
DB 296 CTCTCGAGGAGCTCAGGCAAGCACTGAGGCCCATTTGTAATGATCAGATTCTGAC 355
QY 77 LeuLeuTyrrTyraAaAglAlaLeuAaPaAgaAgluAraAgaSerThraHisAaAgluVal 96
DB 356 TTCCTGCTGATGACGAGGCGCTTGAGCCGAGAGAGACAGACAGATACAGCTACAGTGC 415
QY 97 AlaAlaLeuAaPaAaAaAgluIleleValIgluAgluProValProIleThrlleAgluVal 116
DB 416 ACCCTGAGATGACAGATGACATGCTTGTGGGTCCACACCTGTGCTTGACAGTGC 475
QY 117 LysAaPleAaAaPaAaAaAgluProThraPheLeuAaAgluAaAgluAaAgluAaAgluAaAglu 136
DB 476 AAGGATGAGAAATGACAGGAGTCCCATTTCTCAAGCATTTACAGAGCTGCGCTGAC 535
QY 137 GlinaSerAaAaAaAgluAaAgluAaAgluAaAgluAaAgluAaAgluAaAgluAaAgluAaAglu 156
DB 536 CGGGGTACAGGAGCTGAGTCCCTTCTCTTGAAGGCTTCAAGCCGAGATGAGCA 595
```

QY 157 AlaThrProAsnGlnGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 596 GGCAAGCCCACTCGATCTTCATCTCCATCTCTGAGCCAGGCTCCAGCCCAAGCTTCC 655
 QY 177 ValMetTyrPheGlnIleAsnAsnLeuThrGlyValIleSerLeuThrArgGlnGlySer 196
 DB 656 CCAGACATGTTCCAGCTGAGACCTCGCTGGGGCTCTGGCCCTCCAGCCCAAGGGAGC 715
 QY 197 GlnGlnLeuAsnProAlaValAsnProSerTyrAsnLeuValIleSerValIleAspMet 216
 DB 716 ACCAGCTTGACCAAGCCCTGAGAGACCTTACAGCTGTGGTACAGCTCAAGCATG 775
 QY 217 GlyGlnIleSerGlnAsnSerPheSerAspThrThrSerValIleIleValThrGln 236
 DB 776 GGTATCAGAGGCC--TCAGGCCACAGGCCCTCCAGCTGGAGAGTCTTCATCATAG 832
 QY 237 AsnIleThrValAspProValGlnMetValGlnAsnSerThrAspProHisPro 256
 DB 833 AGCAGCTGGGTGCTCCCTAGAGCTTATCCAGCTGGAGAGATCCAAAGTCTTATCCG 892
 QY 257 IleValIleThrGlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLys 276
 DB 893 CACCAATGGCCCAAGTACACTGGAGTGGGGTATGTGCACTATCACCTG----- 943
 QY 277 GlnLysLeuProAlaGlnPheProPheSerIleAspGlnGlnGlyAspIleTyrValThrGln 296
 DB 944 GAGGAGCATCCCGGGAGCCCTTGAAGTGAATGACAGAGGAACTCTACGTGACAGA 1003
 QY 297 ProLeuAspArgGlnGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGlnTyr 316
 DB 1004 GAGCTGAGACAGAAAGCCAGGCTGAGTACTGCTCCAGGTGGGGCTCAGAAATTCCTAT 1063
 QY 317 GlyLysProLeuSerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsn 336
 DB 1064 GGCGAGGACTATGGGCGCCCTCTGAGCTGACAGCTGGTGAATGAGTGAATGACAAAC 1123
 QY 337 ProProThrCysProSerProValThrValPheGlnValGlnGlnAsnGlnArgLysGly 356
 DB 1124 GTGCTTATCTGCTCCCGCTGACCCCAAGTCAAGTCACTGAGTCACTGACCAAGT 1183
 QY 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGlnGlnAsnThrAlaAsnSerPhe 376
 DB 1184 ACTGAAGTACTAAGCTCTGACAGAGAGATGACAGTCCCGGGCTCCCAATTCCTCAC 1243
 QY 377 LeuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
 DB 1244 GTTGTGATACAGCTCTGAGCCCTGAGCTGAGATGGGATGAGAGGGAGAGCCCTCAG 1303
 QY 396 IleGlnIleThrTyrAlaGlnMetLeuGlnLeuAlaLysGlnSerLeuLysVal--GlnAsp 414
 DB 1304 GTGAGCCCACTTCAAGGAGTGTAGCTGGGGGTCTCCATCTCCAGCAGGCGCAAGAC 1363
 QY 415 ThrProGlnTyrAsnLeuThrIleGlnValSerAspLysAsp-----PheLysThrLeu 432
 DB 1364 ATCTGCTTCTGGTCTGGCCATGAGCTGGCAGCGCAGAGGGGTGCTTACAGACAGC 1423
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnLysSer 452
 DB 1424 TGTGAAGTCAAGTCAAGTCAAGATATCATATCATAGCCCTGAGTTCATCATCTCC 1483
 QY 453 AspTyrGlnLeuLeuThrLeuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1484 CAGATTGGCCCTAAAGCTCTCTGAGATGTGAGCCCGGAGCTCTGAGCCATGCTA 1543
 QY 473 GlnAlaThrAspAlaAsp--GluProPheThrGlySerSerLysIleLeu--TyrHis 490
 DB 1544 ACAAGCCATTGATGCTGACCTGAGCC-----GCTTCCCGCTCATGATTTTCC 1594
 QY 491 IleIleLysGlnAspSerGlnGlyArgLeuGlnValAspThrAspProHisThrAsnThr 510
 DB 1595 ATTGAGAGGGAGACAGAAAGGAGCTTTGGCTCGATTTGGAGGCA-----GACTCT 1648

QY 511 GlyTyrValIleIleLys-----LysProLeuAspPheGlnThrAlaValSerAsn 528
 DB 1649 GGCACTTTAGACTGACAGCTGCAAGAACTCAGTATGAGGCACTCCAGTATGAG 1708
 QY 529 IleValPheLysValGlnAsnProGlnProGlnValPheGlyValLysTyrAsnAlaSer 548
 DB 1709 GTGGTGGTGGTGTGAGAGTGGGAGAGCTGGTGT--GGGCGAGCCCAAGCCCTGAG 1765
 QY 549 SerPheAlaLysPheThrLeuIleValThrAspValAsnGlnAlaProGlnPheSerGln 568
 DB 1766 GCCACCGCCAGGCTGCTGTGTAGTGAAGAGTATGCCACCCCAAGTTCAGCCAG 1825
 QY 569 HisValPheGlnAlaLysValSerGlnAspValAlaIleGlyThrLysValGlnAsnVal 588
 DB 1826 GAGAGCTAGAGGCGAGGTGCTCCCATCAGTGGCCCGGAGCTTCTCCCTGAGCACTC 1885
 QY 589 ThrAlaLysAspProGlnGlnLysAspIleSerTyrSerLeuArgGlnValAspArgGly 608
 DB 1886 CAGCCCTCGAGCCCTTCAAGCCGAGACCTCAGGTTCTCTGATCAATGATCAGAGGC 1945
 QY 609 TrpLeuLysIleAspHisValThrGlnGlnIlePheSerValAlaProLeuAsp--Arg 627
 DB 1946 TGGCTTGATTTGAATTTCTCGGGAGGTGCACACCGCCAGTCTTCAGAGGGCGCC 2005
 QY 628 GlnAlaGlySerProTyrArgValGlnValAlaIleValThrGlnValGlyLysSerLeu 647
 DB 2006 CAGCTGGGAGACCTTACACAGTGTGTGAGAGGCCAGAGTACAGCC----- 2053
 QY 648 SerSerValSerGlnPheHisLeuIleLeuMetAspValAsnAspAsnProAlaGln 667
 DB 2054 -----CTGACTTGGCCCTGCTGCTCCCTCCCA----- 2080
 QY 668 AlaLysAspTyrThrGlnLeuPhePheCysHisProLeuSerAlaProGlySerLeuIle 687
 DB 2081 -----TACTCTGACACCCCGCCAGACCATGCTGATCTG 2119
 QY 688 Phe--GlnAlaThrAspAspArgGlnHisLeuPheArgGlyProHisPheThrPheSer 706
 DB 2120 AGTGAACCAAGACAGAGCCCGCATCTGGCCAGTGGGAGGAGTCC--TACAGCTTACAC 2176
 QY 707 LeuGly--SerGlySerLeuGlnAsnAspTrpGlnValSerLysIleAsnGlyThrHis 725
 DB 2177 CTGTGCTCCCAACCCCAAGGAGTGAAGGATGGCGGCTCCAGACTCTCATATGCTCCAT 2236
 QY 726 AlaArgLeuSerThrArgHisThrAspPheGlnGlnAlaTyrValValIleuLeuArg 745
 DB 2237 GCTTACTTACCTTGGCCCTGCACTGGGTGGAGCCACGTGAACATATCCCGGTG 2296
 QY 746 IleAsnAspGlyLysArgProProLeuGlnGlyIleValSerLeuProValThrPheCys 765
 DB 2297 GTACGCCCAATGCCAG-----ATGTGCACTGCTGTGTAAGTATGATGTGT 2347
 QY 766 SerCys--ValGlnGlySerCysPheArgProAlaGlnHisGlnThrGlyIleProThr 784
 DB 2348 CGCTGCAACGTGAGGGGCGAGTCAATGCCAGGTGGCGCGCATGAAGGCGAGGCCAGC 2407
 QY 785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAla 804
 DB 2408 AAGCTGTGGCAATGGGCATCTTGTAGGCACTGTGTACAAATAGGAATCTTCTCATC 2467
 QY 805 ValValPheIleArgIleLysLysAspLysGlyLysAspAsnValGlnSerAlaGlnAla 824
 DB 2468 CTGATTTTCAACCACTGAGACATGTCAGAGAAAGAGACCGGATCAACAGAGACAGC 2527
 QY 825 SerGlnValLys 828
 DB 2528 GTGCCCCCTGAAG 2539

Search completed: June 22, 2003, 03:29:38
 Job time : 417 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:09:40 ; Search time 114 Seconds
(without alignments)
2238.202 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321
Sequence: 1 MIQAHLSHSLCLMYLATG.....DKGKNVSAQSEVPLNS 832

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop. 6.0 , Delext 7.0	

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xid
-O=/cgsm2_1/USPTO_spool/US10025380/rnat_17062003_173328_24692/app_query.fasta_1.1031
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MNLIN=0 -MAXLEN=200000000
-USER=US10025380 @CGM 1.1.39 @rnat_17062003_173328_24692 -NCPUI=6 -ICPI=3
-NO MAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents NA:*
2: /cgsm2_6/ptodata/2/ina/5A COMB.seq:*
3: /cgsm2_6/ptodata/2/ina/5B COMB.seq:*
4: /cgsm2_6/ptodata/2/ina/6A COMB.seq:*
5: /cgsm2_6/ptodata/2/ina/6B COMB.seq:*
6: /cgsm2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4321	100.0	2499	1	US-08-431-560-2
2	4321	100.0	2499	1	US-08-463-345-2
3	684.5	15.8	3048	1	US-08-188-228-47
4	684.5	15.8	3048	1	US-08-332-643-41
5	684.5	15.8	3048	1	US-08-332-643-41
6	671.5	15.5	3552	4	US-09-643-597-126
7	660.5	15.3	2690	1	US-08-188-228-61
8	660.5	15.3	2690	1	US-08-332-643-55
9	660.5	15.3	2690	1	US-08-332-643-55
10	630.5	14.6	2779	1	US-08-474-067-3
11	630.5	14.6	2779	2	US-08-474-068A-3
12	630.5	14.6	2779	2	US-08-472-481-3

13	630.5	14.6	3959	1	US-08-474-067-1	Sequence 1, Appl1
14	630.5	14.6	3959	2	US-08-472-481-1	Sequence 1, Appl1
15	630.5	14.6	3959	2	US-08-472-481-1	Sequence 1, Appl1
16	593	13.7	2808	1	US-08-237-919-1	Sequence 1, Appl1
17	593	13.7	2808	4	US-08-732-429-1	Sequence 1, Appl1
18	593	13.7	2808	4	US-09-798-267-1	Sequence 1, Appl1
19	593	13.7	2808	5	PCT-US95-05518-1	Sequence 1, Appl1
20	574.5	13.3	3581	2	US-08-738-349-1	Sequence 1, Appl1
21	563	13.0	346	4	US-09-221-298-21	Sequence 21, Appl1
22	562.5	13.0	3712	4	US-08-738-349-3	Sequence 21, Appl1
23	560	13.0	627	4	US-09-385-982-514	Sequence 514, App
24	555.5	12.9	2625	1	US-08-188-228-57	Sequence 57, Appl1
25	555.5	12.9	2625	1	US-08-332-643-51	Sequence 57, Appl1
26	555.5	12.9	2625	1	US-08-332-643-51	Sequence 57, Appl1
27	550.5	12.7	3136	1	US-08-188-228-41	Sequence 41, Appl1
28	550.5	12.7	3136	1	US-08-332-643-41	Sequence 41, Appl1
29	546.5	12.6	2550	1	US-08-188-228-53	Sequence 53, Appl1
30	546.5	12.6	2550	1	US-08-332-643-47	Sequence 47, Appl1
31	546.5	12.6	2550	1	US-08-332-643-53	Sequence 53, Appl1
32	545.5	12.6	3914	2	US-08-738-349-5	Sequence 5, Appl1
33	532	12.3	2521	1	US-08-188-228-59	Sequence 59, Appl1
34	532	12.3	2521	1	US-08-332-643-53	Sequence 59, Appl1
35	532	12.3	2521	1	US-08-332-643-53	Sequence 59, Appl1
36	509.5	11.8	2490	1	US-08-188-228-45	Sequence 45, Appl1
37	509.5	11.8	2490	1	US-08-332-643-45	Sequence 45, Appl1
38	481.5	11.1	3164	1	US-08-188-228-49	Sequence 49, Appl1
39	481.5	11.1	3164	1	US-08-332-643-43	Sequence 43, Appl1
40	481.5	11.1	3164	1	US-08-332-643-49	Sequence 49, Appl1
41	442.5	10.2	4104	1	US-07-998-003A-94	Sequence 94, Appl1
42	442.5	10.2	4104	1	US-08-453-274B-94	Sequence 94, Appl1
43	442.5	10.2	4104	1	US-08-453-695A-94	Sequence 94, Appl1
44	442.5	10.2	4104	1	US-08-268-161A-94	Sequence 94, Appl1
45	442.5	10.2	4104	2	US-08-453-702A-94	Sequence 94, Appl1

ALIGNMENTS

RESULT 1
US-08-431-560-2
Sequence 2, Application US/08431560
Patent No. 5620855
GENERAL INFORMATION:
APPLICANT: Anne H. Dantzig, et al.
TITLE OF INVENTION: Mammalian Influx Peptide
TITLE OF INVENTION: Transporter
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/431,560
FILING DATE: 01-MAY-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/013,462
FILING DATE: 04-FEB-1993
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2499 nucleotides
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULAR TYPE: DNA
US-08-431-560-2

Alignment Scores:

Pred. No.:	0	Length:	2499
Score:	4321.00	Matches:	832
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-025-380-1081 (1-832) x US-08-431-560-2 (1-2499)

```

Qy 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
Db 1 ATGATACTTCAGGCCCACTTCTCACTCCCTGTGTCTTCTTATGCTTATTTGGCACTGGA 60
Qy 21 TyrGlyGlnGlyGlyLysPheSerGlyProLeuLysProMetThrPheSerTyrGlu 40
Db 61 TATGGCCCAAGAGGGGAAGTTTATGTGACCCCTGAAACCATGACATTTTCTATTTATGAA 120
Qy 41 GlyGlnGlyProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValThrPhe 60
Db 121 GGCCCAAGAACGAGTCAATATATTCAGTTTAAAGCCAACTCTCTGCTGTGACTTTT 180
Qy 61 GluLeuThrGlyGlyLysThrAspAsnIlePheValIleGluArgGlyGlyLeuLeuTyrTyr 80
Db 181 GAATTAAGTGGGAGAGACAGACACATATTTGTATAGAACGGAGGGACTTCTGTATATAC 240
Qy 81 AsnAlaGlyAlaLeuAspArgGlyLysThrArgSerThrHisAsnLeuGlnValAlaLeuAsp 100
Db 241 AACAGAGCTTGGACAGGAAACAAGATCTACTACATCTCCAGTTTGCAGCCCTGGAC 300
Qy 101 AlaAsnGlyIleIleValGlnGlyProValProIleThrIleGluValLysAspIleAsn 120
Db 301 GCTATGGAATTAATAGAGAGGAGTCCAGTCCCTATACACATAGAAAGAAAGACATAC 360
Qy 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlyGlySerValArgGlnAsnSerArg 140
Db 361 GACATGACACCAAGTTTCTCCAGTCAAAAGTACGAAGCTCAGTAAAGCAAACTCTCC 420
Qy 141 ProGlyLysProPheLeuTyrValAsnAlaThrAspLeuAspProAlaThrProAsn 160
Db 421 CCAGAAAGCCCTTCTGTATGTCAATGCCAGAGCTGTGATGCCGACCTCCCAAT 480
Qy 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
Db 481 GGCCAGCTTATATCCAGATGTCAATCCAGCTTCCAGATCAACAATGTATGACTTT 540
Qy 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlyGlySerGlnGlyLeuAsn 200
Db 541 CAGATCAACAACAAGGAGGAGCATCTCTTACCGAGAGGATCTCAGAAATTGAT 600
Qy 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
Db 601 CCTCTAAGATCCCTTCTATATCTGTGTGATCTCAGTGAAGGACATGGGAGGCGCAGAT 660
Qy 221 GluAsnSerPheSerArgThrThrSerValAspIleIleValIleThrGlnAsnIleThrLys 240
Db 661 GAGATTCCTTCAGTGAATCAACATCTGTGATATCATAGACAGAAATATTTGGAA 720
Qy 241 AlaProLysProValGlnMetValGlnAsnSerThrAspProHisProIleLysIleThr 260
Db 721 GCACCAAAACCTGTGGAGATGTGGAAAACTCAACTGATCTCCACCCCATCAAAATCACT 780
Qy 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGlyLysLeuPro 280
Db 781 CAGGTGGCGTGAATGATCCCGGTGCCAATATTTCTTAACTTGAACAAGAAACCTGCCA 840
Qy 281 ArgPheProPheSerIleAspGlnGlyAspIleTyrValIleThrGlnProLeuAspArg 300
Db 841 AGATTTCCATTTCAATTGACCAAGAGAGATATTTACGTGACTCAGCCCTTGACCGA 900
Qy 301 GlyGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGlyTyrGlyLysProLeu 320

```

```

Db 901 GAAGAAAAGAGTCATATGTTTTTTATGCAAGTTCGAAGAGTACGGAACAACACTT 960
Qy 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
Db 961 TCATATCCGTGGAAATTCATGTAAAGATTAAGATTTATATATATATCCACTACATGT 1020
Qy 341 ProSerProValThrValPheGluValGlnGluAsnGlyLysArgLeuGlyLysSerIleGly 360
Db 1021 CCGTCAACAGTAAACCGATTTAGAGTCCAGAGAAATGAACAGATCGGGTAACTATCCGG 1080
Qy 361 ThrLeuThrAlaHisAspArgAspGlyGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
Db 1081 ACCCTTACTGCATATGACAGAGGATGAAGAAATACCTGCCAAGCTTTCTAACTACAG 1140
Qy 381 IleValGlnGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
Db 1141 ATGTGGAGAACTCCCAAACTCCCAATGAGAGACTCTTCTTAATCCAAACCTATGCT 1200
Qy 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrProGlnTyrAsnLeu 420
Db 1201 GGAATGTACAGTTAGCTAAACAGTCTTGAGAGAACAGATACCTCTGACTCAACTTA 1260
Qy 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
Db 1261 ACGATAGAGGTGTCTGACAAAGATTTCAAGACCCCTTGTGTGCAAAATCAACGTTATT 1320
Qy 441 AspIleAsnAspGlnIleProIlePheGlyLysSerAspTyrGlyAsnLeuThrLeuAla 460
Db 1321 GATATCAATATATGATCCCATCTTTGAAAATTCAGATTTATGAAACCTGACTCTTGCT 1380
Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
Db 1381 GAAGACACAAACATTTGGTCCACCATTTAACAATCCAGGCCACTAGCTGATGAGCCA 1440
Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleGlyLysAspSerGlyLysArgLeu 500
Db 1441 TTTACTGGAGATTCTAAATTTCTGTATCATATCATAAAGGAGACAGTGAAGCGCTG 1500
Qy 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
Db 1501 GGAGGTGACACATATCCCATACCAACACCGATATGTCTAATTAATAAGGCTCTTGAT 1560
Qy 521 PheGlyLysThrAlaValSerAsnIleValPheLysAlaGlyLysProGluProLeuVal 540
Db 1561 TTTGAAACAGCACTGTTTCCACATTTGTTCAAGAGAAATCTGAGCCTCAAGT 1620
Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
Db 1621 TTTGTGTGAAGTACATGACATGCTTTTGGCCAAAGTTCACGCTTATTTGACAGATGTG 1680
Qy 561 AsnGlyLysAspProGlnPheSerGlnHisValPheGlnAlaLysValSerGlyAspValAla 580
Db 1681 AATGAAGCACCTAATTTTCCCAACAGTATTCAGAGCAAAATCAGTGAAGTGAAGT 1740
Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGlyLysLeuAspIleSerTyr 600
Db 1741 ATAGGCACTTAATGGGCAATGTGACTGCCAAGATCCGAAGAGTCTGGACATTAAGTAT 1800
Qy 601 SerLeuArgGlyAspThrArgGlyTyrPLeuLysIleAspHisValThrGlyLysIlePhe 620
Db 1801 TCACTGAGGGAGACACAAGAGGTTGGCTTAAATTTGACACAGCTGTGTGAGATCTTT 1860
Qy 621 SerValAlaProLeuAspArgGlyAlaGlySerProTyrArgValGlnValValAlaThr 640
Db 1861 AGTGTGCTTCATTTGACAGAGAGCGGAAGTCATATCGGTATACAGTGGGCGCAC 1920
Qy 641 GluValGlyLysSerLeuSerSerValSerGlyPheHisLeuIleLeuMetAspVal 660
Db 1921 GAAGTAGGGGGGCTTCTTAACTCTGTCTGAGATTTCAACTGATCTTATGAGATGTG 1980
Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPheCysHisProLeu 680
Db 1981 AATGACAAACCTCCAGGCTAGCCAAAGACTACACGGGCTTGTCTTCTGCAATCCCTTC 2040

```

QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
 Db 2041 AGTCACTCTGGAAGTCTCATTTTCGAGGCTACGATGATGATGACGACATTTTCGGGGT 2100
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnProGluValSerLys 720
 Db 2101 CCCCATTTTACATTTTCCCTCCGGAGGAGGAGCTTACAAACGACTCGGAGAGTTTCCAAA 2160
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisPhePheGluGluArgAlaTyr 740
 Db 2161 ATCAATGATGACTATGCCCCGACCTGTCTTACAGGACACAGACTTTTGAGAGAGGCGCTAT 2220
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu 760
 Db 2221 GTGCTTGTGATCCGCATCAATGATGAGGGGTCGGCACCTTGGAAGGCAATGTTTCTTTA 2280
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2281 CCAATTACATTTCTGCAGTTGTGTGAGAGAGAGTTGTTCCGGCCAGCAGGTCAACGACT 2340
 QY 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2341 GGGATACCCGACTGTGGGACATGGGAGTGATGATGCTGACACCTTCTGTGTGATTTGT 2400
 QY 801 IleIleLeuAlaValAlaPheIleArgIleLysIleAspLysGlyLysAspAsnValGlu 820
 Db 2401 ATATATTTTTCAGTGTGTCTTATCCGCATTAAGAGAGATTAAGGCAAAAGTATTTTGAA 2460
 QY 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2461 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAAAGC 2496

RESULT 2

US-08-463-345-2
 ; Sequence 2, Application US/08463345
 ; Patent No. 5710018
 ; GENERAL INFORMATION:
 ; APPLICANT: Anne H. Pantzig, et al.
 ; TITLE OF INVENTION: Mammalian Influx Peptide
 ; TITLE OF INVENTION: Transporter
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Eli Lilly and Company
 ; STREET: Lilly Corporate Center
 ; CITY: Indianapolis
 ; STATE: IN
 ; COUNTRY: U.S.A.
 ; ZIP: 46285
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
 ; COMPUTER: Macintosh
 ; OPERATING SYSTEM: Macintosh
 ; SOFTWARE: Microsoft Word
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/463,345
 ; FILING DATE: 05-JUN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/013,462
 ; FILING DATE: 04-FEB-1993
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2499 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; US-08-463-345-2

Alignment Scores:

Score: 0 Length: 2499
 Pred. No.: 4321.00 Matches: 832

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 1
 Gaps: 0

US-10-025-380-1081 (1-832) x US-08-463-345-2 (1-2499)

QY 1 MetIleLeuGlnAlaHisIleuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 1 ATATATCTTCAAGGCCCATCTCTCCCTGTGTCTTATGTTATTGGCACTGGA 60
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 Db 61 TATGGCCAAGAGGGGAAAGTTAGTGACCCCTGAACCAATGACATTTCTATTATTAAGA 120
 QY 41 GlyIleGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPhe 60
 Db 121 GGCCAAAGAACCGGTAAATTAATTAATTCAGTTTAAAGCCCAATCCTCTGTGACTTTT 180
 QY 61 GluLeuThrGlyGluThrAspAsnIlePheValIleGluArgGlyLeuLeuTyrTyr 80
 Db 181 GAATTAATCTGGGAGACAGACAAATATTGTTGATGAACGGGAGGACTTCTGTATTAC 240
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 241 AACAGAGCCTTGACAGGGAACAAAGATCTACTCAATCTCCAGGTGACGCTTGAC 300
 QY 101 AlaAsnGlyIleIleValAlaGlyProValProIleThrIleGluValLysAspIleAsn 120
 Db 301 GCTAATGGAATTAATTAATGAGGGGTCCAGTCCCTATCAATGAATGAAGGACATCAAC 360
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg 140
 Db 361 GACAAATGACCCGACTTTTCTCCAGTCAAAAGTACGAAGGCTCAGTAAGGACGAATCTCCG 420
 QY 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspAspProAlaThrProAsn 160
 Db 421 CCAGGAAGAGCCCTTCTTGTATGTCATGACAGACAGACTGGATGATCCGGCACTCCCAAT 480
 QY 161 GlyIleLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 Db 481 GGCCAGCTTATTAACAGATGTCATCCAGCTTCCATGATCAATGATGATGATGATCTT 540
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 Db 541 CAGATTAACAACAAACGGGAGGCAATCTCTTACCCGAGAGGATCTCAGGAATGAAT 600
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 Db 601 CCGCTTAAGATCTCTCTATATATCGTGATCTCATGAAGACATGGGAGCCAGAGT 660
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
 Db 661 GGAATTCCTTCAGTATACCAATCTGCGATATCATATGATGACAGAGATATTGGAAA 720
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 Db 721 GACCAAAACCTGTGAGATGAGTGAAGAACTCAACGATCTCACCCATCAAAATCACT 780
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 Db 781 CAGGTGCGGTGAGATATCTCCGTGACAAATATCTTGTGTTGACAAAGAGAGCTGCCA 840
 QY 281 ArgPheProPheSerIleAspGlnGluLysAspIleTyrValThrGlnProLeuAspArg 300
 Db 841 AGATTCCCAATTTTCATTTACCAAGAGAGAGATATTATGATCACTCAGCCCTTGACCCA 900
 QY 301 GluGluLysAspAlaThrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 Db 901 GAAGAAAGAGATGATATGTTTATGCAAGTTGCAAGAGATGATGAGAAACCACTT 960
 QY 321 SerTyrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrCys 340

Db 961 TCATATCCGCTGGAATTCATGTAAGATTAAAGATTAATGATTAATCAACCTACTACTGT 1020
 Qy 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db 1021 CCGTCACCAAGTAAACGATATTGAGTCCAGAGATTAACAGACTGGGTAAACAGATCGGG 1080
 Qy 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db 1081 ACCCTTACTGCACTGACAGGAGTAAAGAAATCTCCCAACGTTTCTTAACTCAAGG 1140
 Qy 381 IleValGluGlnThrProTyrLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db 1141 ATGTGGAGCAAACTCCCAAACTTCCCATGATGATGATCTTCCAAATCCAAACCTTAATGT 1200
 Qy 401 GlyMetLeuGlnLeuAlaValArgGlnSerLeuIleValGlnAspThrProGlnTyrAsnLeu 420
 Db 1201 GGAATGTTACAGTTAGCTTAAACAGTCTCTGAAGAGCAAGTACTCTCTGAGTCAACTTA 1260
 Qy 421 ThrIleGluValSerAspLysAspPheLysPheLysPheValGlnIleAsnValIle 440
 Db 1261 ACGATTAAGGTGTGTGACAAAGATTTCAAGACCTTTGTGTTTGTGCAAAATCAAGTTAT 1320
 Qy 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db 1321 GATATCAATGATGATCCCATCTTGAATAATCAGATTATGGAATCACTGACTTGTCT 1380
 Qy 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 Db 1381 GAAGACCAACAACTGGGTGTCCACATCTTAAACCTCAAGGCACTGATGCGATGAACCA 1440
 Qy 481 PheThrGlySerSerLysIleLeuTyrHisIleIleLeuGlyAspSerGluGlyArgLeu 500
 Db 1441 TTTTACTGGAGCTTAAATTTCTGTATCATATCTAAAGAGACAGTGAAGGGAGCGCTG 1500
 Qy 501 GlyValAspThrAspProHisAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 Db 1501 GGGGTGACACAGATCCCATACCAACACCGATATGCTAATTAATAAAGCCCTTGAT 1560
 Qy 521 PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal 540
 Db 1561 TTGGAACAGACGCTGTTCCACATGTTGTTCAAGACAGAAATCTCGAACCTCTAGTG 1620
 Qy 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 Db 1621 TTTGTGTGAGTACATGCAATGCAATGTTTTCCTTTCACAGTTCAGCTTATTTGACAGATG 1680
 Qy 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 Db 1681 AATGAACACACTCAATTTCCCAACAGTATTCAGAGAAAGTCAAGTGAAGATGTAGCT 1740
 Qy 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTyr 600
 Db 1741 ATAGGCACTAAAGGCGCAATGTACTGCCAAGATCCAGAAAGTCTGGAATAGATGAT 1800
 Qy 601 SerLeuArgLysAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
 Db 1801 TCATCTAGGGAGACACAGAGGTTGGCTTAAATTTGACACGCGACTGGGAGATCTTT 1860
 Qy 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 Db 1861 AGTGTGCTCATTTGACAGAGAGAGCGGAGTCCATTCGGGTACAAAGTGTGGCCACA 1920
 Qy 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db 1921 GAACTAGGGGGGTCTTCTTAAAGCTGTGTCAAGTGTCCACGATCTCTTATGATGATG 1980
 Qy 661 AsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 1981 AATGACAACTCCCAAGCTAGCCAGACACAGACTACAGGGCTTCTTCTGCAATCCCTC 2040
 Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 2041 AGTCACCTGGAAGTCTCATTTTGAAGGCTACTGATGATGATGATGATGATGATGATGAT 2100

Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerLys 720
 Db 2101 CCCCATTTTACATTTTCCCTCGCAGTGAAGCTTACAAAGACAGCTGGAAAGTTTCCAA 2160
 Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db 2161 ATCAATGATCATCATCCCGACACTGTCTTACAGCACACAGACTCTTGAGAGAGGGGTAT 2220
 Qy 741 ValValLeuIleAlaArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db 2221 GTCTGTTAATCCGATCAATGATGAGGGGTCGGCCACCTTGGAGAGCATTTGTTTCTTA 2280
 Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 Db 2281 CCAATTACATCTTCGACGTTGTGTGAGAGAAAGTGTTCGGGCACAGCTCAACAGACT 2340
 Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2341 GGGATACCACTGTGGGAGTGGCAGTGGTATCTGACCAACCTTCTGTGATTTGT 2400
 Qy 801 IleIleLeuAlaValAlaPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
 Db 2401 ATAAATTTAGCACTTGTGTTATCCCATTAAGAGAAATTAAGCAAGATTAATGTTGAA 2460
 Qy 821 SerAlaGlnAlaSerGluValLysProLeuArgSer 832
 Db 2461 AGTCCTCAAGCATCTGAAGTCAAACTCTGAAGAC 2496

RESULT 3
 US-08-188-228-47
 ; Sequence 47, Application US/08188228
 ; Patent No. 5597725
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSER: Borun
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/188,228
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/049,460
 ; FILING DATE: 19 APR 1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/872,643
 ; FILING DATE: 17 APR 1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5597725and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 47:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3048 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single


```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-188-228-47

Alignment Scores:
Pred. No.: 4,4e-65      Length: 3048
Score: 684.50          Matches: 176
Percent Similarity: 46.92%      Conservative: 106
Best Local Similarity: 29.28%    Mismatches: 288
Query Match: 15.84%            Indels: 31
DB: 1                      Gaps: 12

US-10-025-380-1081 (1-832) x US-08-188-228-47 (1-3048)

QY      239  TrrpysalAprolysProvalGluMetValGluasnSerThrAspProHisProIleLys 258
      525  TGGGTATGCCA---CCCATCAAGTGGCCGAGAACTCGCGCGGCGCTTCCCGGCACAG 581
QY      259  IleThrGlnValAArgTrp-----AsnAspProGlyAlaGlnTyrSerLeu----- 273
      582  CTCGTGAGGATCCCGTCCGACAAAGACATGATCCCATCCGCTACAGATCAGCGGA 641
QY      274  ValAspArgGluLysLeuProArgPheProPheSerIleAsp---GlnGluGlyAspIle 292
      642  GTGGGTCCGACCGACCGCCCGCATGAGAGTCTTCAGCATTAATCTCATGTCGCGCGGATG 701
QY      293  TyrValThrGlnProLeuAspArgGluGlnLysAspAlaTyrValPheTyrAlaValAla 312
      702  TAGGTCAAGAGCCCATGAGACCGGAGAGACCCCTTACCATCCATCCGACCGCACGCT 761
QY      313  LysAspGluTyrGlyLysProLeuSerTyrProLeuGluIleHisValLysValLysAsp 332
      762  GTGACATGATGATGCAACAGGTGAGAGAACCCCATCACTGATCATCATCATGATGAC 821
QY      333  IleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGluAsn 352
      822  ATGATATACAAACCACTGAGTTTCATCAACAGGTGTTCACATGCTCGGTGAGACGAGGC 881
QY      353  GluArgLeuGlyAsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThr 372
      882  TTCAGGCGACGACCTACGTGATGATGATCCATCAGGCCCAACGATGCTGACGACAGACAG 941
QY      373  AlaAsnSerPheLeuAsnTyrArgIleValGlnGlnThrProIlyLysLeuProMetAspGly 392
      942  GCCAAGCGAGTGGCGGTACCGGATCGTGACCCAGACCCCAAGACCGGTGCCAAGAT 1001
QY      393  LeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys 412
      1002  ATGTTCACCATCAACAGCAGAGATGAGATATGCTCACAGTGGCGGCTGGTGGACCGA 1061
QY      413  GlnAspThrProGlnTyrAsnLeuThrIleGluValSerAspLysAsp----- 428
      1062  GAGAAAGTTACAGATGACAGATCATGTTCAAGGCCACAGATATGAGAAAGAAATCTTAC 1121
QY      429  -----PheLysThrLeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIle 446
      1122  TATGGCTCTCAAAACACAGCCACAGCATCATCAGGTGACAGATGAGATGACAAACCG 1181
QY      447  ProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGly 466
      1182  TCAGAAATTACCGCCAGCAGCTTTCAGAGG-----GAGGTCCCGCAAAACAGGTGTGAG 1235
QY      467  SerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPheThrGlySerSerLys 486
      1236  ACCGTGTCCCAAAACCTCAGCGTGAATGACCGAGATACGCCCATCTCCAAACTGGAAT 1295
QY      487  IleLeuTyrHisIleIleLysGlyAspSerGluGlyArgLeuGlyValLysAspThrAspPro 506
      1296  GCCGTTCACCGCATCATCACTGGGAGATCCATCCGGGACATTCAGCGCCACAGACCCC 1355
QY      507  HisThrAspThrGlyTyrValIleIleLysLysProLeuAspPheGluThrAlaIleVal 526
      1356  GTAAACCAACAGGCGATGTCAACCGTGTGAAGCAGATGACTGACGAGTCAACAGAGCT 1415

```

```

QY      527  SerAsnIleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrAsn 546
      1416  TTCATGCTGACAGATGATGTTGCCAACAGAGCCCTTGGCCAGCGGATATTCAGATCTCC 1475
QY      547  AlaSerSerPheAlaLysPheThrLeuIleValThrAspValGlnGluAlaProGlnPhe 566
      1476  TTCAGTCCAGCGCAGGAGGTGACCATCTCCATGATGACATGACATCAACAGAGGCTCTTCT 1535
QY      567  SerGlnHisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysValGly 586
      1536  CCTCAAAACCAAGCTGATCCGCTGAGAGAGGGGTGCCCCCGGACCGTGTACC 1595
QY      587  AsnValThrAlaLysAspProGluGlyLeu-----AspIleSerTyrSerLeuArg 603
      1596  ACCTTTACAGCTGTGAGACCTTACCGGTTTATCATGACAGAGGCTGTGATATCTAAAGCTG 1655
QY      604  GlyAspThrArgGlyTyrLeuLysIleAspHisValThrGlyLysIlePheSerValAla 623
      1656  TCAGACCCAGCGAGCTGTGACATCAATGACCAACAGCGCAGATCAACAGCGTGGCA 1715
QY      624  ProLeuAspArgGluAla-----GlySerProTyrArgValGlnValAla 639
      1716  GTCTGACCGGTGATGCTCTTACACCAAAACAGCTTACGAGGCCACTTCTGGCA 1775
QY      640  ThrGluValGlyLysSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAsp 659
      1776  GCTGACAAATGGATATACCCCGGACGGAGCGGACCGGACCTCCAGATCTATCATTTGAC 1835
QY      660  ValAsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePheCysHisPro 679
      1836  ATCAACGACAAACCCCTTGAGTGGTCCCAAGAGAGCGCAGATGTCGAGAGGCCAAC 1895
QY      680  LeuSerAlaProGlySerLeuIlePheGlnAlaThrAspAspArgGlnHisLeuPheArg 699
      1896  CTGACGCGC-----ATCAACATCAGCGCGCGGACGTGACGTGACGCCAACATC 1946
QY      700  GlyProHis---PheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluVal 718
      1947  GGCCTTACGCTTTCAGAGTCCGCTTTCGCCGCGCGCGCGTGGGAGAAATGACATC 2006
QY      719  SerLysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArg 728
      2007  ACCGCTTGAACGATGATGATGCCAACTCAGCTTGCAGTCTGATCGTGAAGGCGGG 2066
QY      739  AlaTyrValValLeuIleArgIleAsnAspGlyValArgProProLeuGlnGlyIleVal 758
      2067  ATGTATGACGTCCCATCATCTGTCACAGACTCTGGAACCTCTCCCTGTCACACAGTCC 2126
QY      759  SerLeuProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHis 778
      2127  ATATCAAGTCAAGGTGTGCCATGTATGACACAGGGGACGTGCACACCATTTGC--- 2183
QY      779  GlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuVal 798
      2184  -----GCAGTGGACAGCGGTGTGGTGGGACCGGATGTCGATTCATCTCATCTGCG 2237
QY      799  IleGlyIleIleLeuAlaValIle-----PheIleArgIleLysLysAspLysGlyLys 816
      2238  ATCTCATCTCTGTCACATGCTCTGCTGTGTGATGATGATGAAGGCGGAGAGAG 2297
QY      817  Asp 817
      2298  GAG 2300

```

RESULT 4
US-08-332-643-41

; Sequence 41, Application us/08332643

; Patent No. 5639634

; GENERAL INFORMATION:

; APPLICANT: Suzuki, Shintaro

; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

; NUMBER OF SEQUENCES: 56

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ADDRESSEE: Bicknell
 STREET: Two First National Plaza, 20 South Clark
 STREET: Street
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60603
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,643
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/872,643
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5639634and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/30795
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 346-5750
 TELEFAX: (312) 984-9740
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3048 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-332-643-41
 Alignment Scores:
 Pred. No.: 4.4e-65 Length: 3048
 Score: 684.50 Matches: 176
 Percent Similarity: 46.92% Conservative: 106
 Best Local Similarity: 29.28% Mismatches: 288
 Query Match: 15.84% Indels: 31
 Gaps: 12
 US-10-025-380-1081 (1-832) x US-08-332-643-41 (1-3048)
 QY 239 TrpysalAProlyspProValGluMetValGluAsnSerThrAspProHisProIleuys 258
 Db TGGGTATCCCA---CCCATCAACGTGCCCGAGAACTGGCGGGCCCTTCCCGCAGAG 581
 QY 259 IleThrGlnValArgTrp-----AsnAspProGlnAlaGlnTyrSerLeu----- 273
 Db CTCGTGAGGATCCGCTCCGACAAAGACATGATCCCATCCGATACAGATCAGCGAG 641
 QY 582 ValAspIlysgIuylsleuProArgPheProPheSerIleAsp---GlnGluGlyAspIle 292
 Db GTGGGTGCGGACGAGCCCGCCATGAGAGTCTTACGATTAATCCATGATCCGCGGAGTG 701
 QY 274 TyrValThrGlnProLeuAspArgGluGluIlyAspAlaTyrValPheTyrAlaValAla 312
 Db TACGTCAACAGGCCCATGATGACCGGAGAGACAGCGCTTTACACCATCCGAGCCCACT 761
 QY 313 LysAspGluTyrGlyIlyysProLeuSerTyrProLeuGluIleHisValIlyValIlyAsp 332
 Db GTGACATGATGAGCAACAGATGAGAGAACCCCATGACATGATCATGATCATGATGATG 821
 QY 333 IleAsnAspAsnProProThrCysProSerProValThrValPheGluValGlnGluAsn 352
 Db ATGATGATGACCAACCCCTGAGTTCATCAACAGGCTTCAACATGCTCCGTGAGACGAGGC 881
 QY 353 GluArgIleuGlyAsnSerIleGlyThrIleuThrAlaHisAspArgAspGluGluAsnThr 372

Db 882 TCAGAGCAGGACCACTACGATGATCATCATGACGCGCAACGATCTGACAGACACGACG 941
 QY 373 AlaAsnSerPheLeuAsnTyrArgIleValGluGlnThrProIlyAspProMetAspGly 392
 Db GCCACGAGGATGGTGGTATCCGATGTAACCCAGACCCACAGAGCCCGCTCCAGAAAT 1001
 QY 393 LeuPheLeuIleGlnThrTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuIlyys 412
 Db ATGTTCAACCTCAACACGACGAGTATGTCACAGTGGCGGCTGGTGGACCGA 1061
 QY 413 GlnAspThrProGlnTyrAsnLeuThrIleGlnValSerAspIlyAsp----- 428
 Db GAGAAAGTTGACAGATACACATGATCTTCAGGCCACAGATATGAGAGAAATCTCAAC 1121
 QY 429 -----PheIlyThrIleuysPheValGlnIleAsnValIleAspIleAsnAspGlnIle 446
 Db TATGGCTCTCAAAACACGACGACGACGATCATCAGGTGACAGATGATGATGATGATG 1181
 QY 447 ProIlePheGluIlySerAspTyrGlyAsnLeuThrIleuAlaGluAspThrAsnIleGly 466
 Db TCAGATTTTACCGCCAGCAGCTTTCAGAGG-----GAGTCCCGCAAAACAGCTGGAG 1235
 QY 467 SerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPheThrGlySerSerIys 486
 Db ACCGTGGTGCACAACTCAACGATGATGACGAGATCAGCCCACTCTCCAACTGGAAAT 1295
 QY 487 IleLeuTyrHisIleIleIlysgIlyAspSerGlnGlyArgLeuGlyValAspThrAspPro 506
 Db GCCGTTTACGATCATCATGATGAGGATCCATCCGGGCACTTCAAGCTCCGACAGACCC 1355
 QY 507 HisThrAsnThrGlyTyrValIleIleIlyysPheProLeuAspPheGluThrAlaAlaVal 526
 Db GTAACCAACAGGACGATGATCAGCGTGTCAAGGACTGATCAACGATCAACAGAGCT 1415
 QY 527 SerAsnIleValPheIlyAspIlyAsnProGluProIleuValPheGlyValIlyTyrAsn 546
 Db TTCATGCTGACAGTATGATGATGATCCACAGGCGCCCTGGCCAGCGAAATCCAGATCTCC 1475
 QY 547 AlaSerSerPheAlaIlyAspPheThrIleuIleValThrAspValAsnGluIlyAspGlnPhe 566
 Db ATTCAGTCCACGCGGACGAGGATGATCCATTCATGATGATGATGATGATGATGATGAT 1535
 QY 567 SerGlnHisValPheGlnAlaIlyValSerGlnAspValAlaIleGlyThrIlyValGly 586
 Db CCTCAACCAACAGCTGATCCCGCTGAGAGAGGCGCTGCCCGCCGACCGTGTACCC 1595
 QY 587 AsnValThrAlaIlyAspProGluIlyLeu-----AspIleSerTyrSerLeuArg 603
 Db ACGTTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
 QY 604 GlyAspThrArgIlyThrIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 623
 Db TCAGACCCACGAGCTGGCTGACATCAATCAAGCCACCAACGCGGACAGTACACAGGTGGA 1715
 QY 624 ProLeuAspArgGluAla-----GlySerProTyrArgValGlnValAlaIle 639
 Db GTGCTGACCGTATGATCCCTTCAACCAAAACAGCTTCAACAGGCGACCTTCTGTGCA 1775
 QY 640 ThrGlnValGlyIlySerSerIleuSerSerValSerGlnIlyIleuIleuMetAsp 659
 Db GGTGACAAATGATATACCCCGGACGAGGACCGGAGCCCTCCAGATCTATCTCATGATG 1835
 QY 660 ValAsnAspAsnProProAlaGluAlaIlyAspTyrThrIlyLeuPheIlyCysHisPro 679
 Db ATCAACGACCAACCCCTGATGCTGCTGCCAAGAGGCGGATGCTGCGAGAGGCCCAAC 1895
 QY 680 LeuSerAlaProGlySerIleuIlePheGluAlaThrAspAspAspGlnHisIleuPheArg 699
 Db CTGACGCC-----ATCAACATCAACGCGCGGAGCGGCTTACGTCACCCCAACATC 1946
 QY 700 GlyProHis-----PheThrPheSerIleuGlySerIleuGlnAsnAspTyrGluVal 718

Db 1947 GGCCCTACGCTTCGAGCTGCTTTGTCGCCGGCGCCGCGGAGAACTGACCATC 2006
 QY 719 Serlyslleanglythrhlealargleuserthrhghisthrasphelugluarg 738
 Db 2007 ACCCGCTGAACGGTGAATGATGCCCACTGAGCTTGCGCATCTGACGAGCCGG 2066
 QY 739 AlATyValValleulleatgyleasnapslyglargpropleugluglyleval 758
 Db 2067 ATGATGACGCTCCCATCATGCTCAGACACTCTGGAACCTCCCTGCAACAGCTCC 2126
 QY 759 Serleuprovalthrhiecysercyvalgluglysercyshieagprolaelyhis 778
 Db 2127 ATCTCAAGTCAAGTGTGCTCCATGTGATGACAAACGGGAGCTGCACCATTTGGC- 2183
 QY 779 GlnthrlyleleprothrhvalgllymetalaValgllyleleuleuthrhleuleVal 798
 Db 2184 -----GCAGTGGAGGGCGCTGCTGCGGACCGGTGCATCGTGGCATCTCATCTGC 2237
 QY 799 lileglylleleleualavalval-----Pheillearglylelyslasphlylys 816
 Db 2238 ATCTCATCTGCTGACCATGCTCTGCTGTTGTGATGTGATGAGCGGCGAGAGAG 2297
 QY 817 Asp 817
 Db 2298 GAG 2300

RESULT 5

US-08-332-638-47
 Sequence 47: Application US/08332638
 Patent No. 5646250
 GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 STREET: 6300 Sears Tower, 233 S. Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,638
 FILING DATE: 01-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/872,643
 FILING DATE: 17 APR 1992
 APPLICATION NUMBER: US/08/049,460
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5646250and, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31340
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3048 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-332-638-47

Alignment Scores:

Pred. No.: 4,4e-65 Length: 3048
 Score: 684.50 Matches: 176
 Percent Similarity: 46.92% Conservative: 106
 Best Local Similarity: 29.28% Mismatches: 288
 Query Match: 15.84% Indels: 31
 DB: 1 Gaps: 12

US-10-025-380-1081 (1-832) x US-08-332-638-47 (1-3048)

QY 239 TTPlyAlaProlyspProValglumetValgluasnerThrhAspProhispProlelyls 258
 Db 525 TGGGTCAATCCA---CCCATCAAGTGCACGAATACTCGGCGGCGCTTCCGACAG 581
 QY 259 lilethrhValargTTP-----AsnAspProglyalaglntyrserleu----- 273
 Db 582 CTGTGAGATCCCGGTCCGACCAAGACATGATCCCATCCGGTACAGATCAGCGGA 641
 QY 274 ValAspLyglulysleupProArgPheProPheSerlleAsp---GlnluglyAspIle 292
 Db 642 GTGGGTGCGCAGCAGCGCCCATGAGAGGTCTTCAAGATTATCATGTCGCGCGGATG 701
 QY 293 TyrValThrhGlnProleuAspArggluglyllysAspAlaTyrValPheTyrAlaValAla 312
 Db 702 TACGTCAAGAGGCCCATGAGCCGAGAGACAGCCCTTATCCACCTCCAGGCCAGCT 761
 QY 313 LyAspGluTyrGlyslsProleuserTyrProleugluThrhisValylsValylsAsp 332
 Db 762 GTGACATGATATGCAACAAAGTGAAGAACCCCATGACCTGTACATCTTCATATGAC 821
 QY 333 lileAsnAspAspProProthrhCysProSerProValThrhValPhegluValgluasn 352
 Db 822 ATGAATGACAAACACCTGATTCATCAACAGAGTCTCAACAAGTCTCCGTGAGAGAGGC 881
 QY 353 GluArgleuglyAsnserlleGlyThrhleuthrhAlahisAspArggluasnthr 372
 Db 882 TCCAGCGAGGCACTTACGTGATGATGATCATGACCGCCACATGATCTGACGACACACG 941
 QY 373 AlaAsnSerPheleuAsnTyrArglleValgluglyThrhProlyslsleupPrometAspGly 392
 Db 942 GCCAAGCGGATGTGCGGTACCGGATCGTACCCAGACCCCAAGACCCGCTCCAGAAAT 1001
 QY 393 leuPheleulleGlnthrhTyrAlaglymetleuglneulaleuAlasglinsertleuyls 412
 Db 1002 ATTTTACATCAACAGCAGCAGAGTGAATATGTCAAGTGGCGGTGCTGAGACCA 1061
 QY 413 GlnAspThrhProGlnTyrAsnleuthrhleGluValSerAspIysAsp----- 428
 Db 1062 GAGAAAGTTCAAGCAGTACAGATCATGCTTCAAGCCACAGATATGAAAGAAATCTCAAC 1121
 QY 429 -----PheIysrhreucysrPheValGlnlleAsnVallleAspIleAsnAspGlnlle 446
 Db 1122 TATGCGCTTCNAACAGCCAGCCAGCATCATCAGGTGACATGATGAAATGACAAACCG 1181
 QY 447 ProIlePheGluylsSerAspTyrGlyAsnleuthrhleuAlagluAspThrhAsnIleGly 466
 Db 1182 TCAGAAATTTACCGCAGCAGCGTTGAGGG-----GAGGTCCCGAAGAACAGCGTGAAG 1235
 QY 467 SerThrhIleleuthrhleGlnAlaThrhAspAlaAspGluProPheThrhGlySerIys 486
 Db 1236 ACCGTGTGCAAACTCAAGGTGATGAGCCGAGATCAGCCCACTCCAACTGGAAT 1295
 QY 487 lileuthrhIstlellelysglyAspsersgluglyArgleuglyValAspTyrAspPro 506
 Db 1296 GCCGTTTACCGATCATCATGATGATCCATCCGGGCACTTCAAGCTTCGACAGACCC 1355
 QY 507 HisThrhAsnthrhGlyTyrValllelelyslsPProleuAspPhegluThrhAlaVal 526
 Db 1356 GTAACCAAGAGGAGGATGTGACCCGTGTGAAGGAGTGTACAGACTCAACAGAGCT 1415
 QY 527 SerAsnIleValPheylsAlaGluAsnProGluProleuValPhegllyValylsTyrAsn 546
 Db 1416 TTCATGTGACAGTGTATGTATCAACAGGCGCCCTGCGCCAGCGAAATCAGATGTC 1475

Qy	547	AlaseSerPheAlaLysPheThrLeuIleValThrAspValAsnGluAlaProIlePhe	566
Db	1476	TTCCAGTCCACGGCAGGGGTGACATCTCCATCCATGACATGACATGACAGGGCTCTTAATTC	1533
Qy	567	SerGlnHisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysValGly	586
Db	1536	CCCTCAAAACCAACACTGATCGGGCTGGAGAGAGGGCGTGGCCCCCGGACCGGTGTGAC	1599
Qy	587	AsnValThrAlaLysAspProGluGlyLeu-----AspIleSerTyrSerLeuArg	603
Db	1596	ACGTTTTCAGCTGTGGACCTTACACCGGTTTCATGACAGACGGCTGTAAGATATCTCAAGCTG	1655
Qy	604	GlyAspThrArgGlyTyrIleuLysIleAspHisValThrGluAlaPheSerValAla	623
Db	1656	TCAAGCCACGAGAGCTGGCTGCACTCAATGACACCAAGCGCAATGACACCGGTGGCA	1711
Qy	624	ProLeuAspArgGluAla-----GlySerProTyrArgValGlnValAla	639
Db	1716	GTGCTGACACCGGAGTCCCTTACACCAAAACAGTCTAGAGGCCACTTCTCTGGCA	1775
Qy	640	ThrGluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAsp	659
Db	1776	GCTGACAAATGGATATACCCCGGCGACGGCGACCGGAGCTTCAGATCTATCTCATTTGAC	1833
Qy	660	ValAsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLysPhePheCysHisPro	679
Db	1836	ATCAACGACACAGCCCTCTGACGTGCTGCCCAAGAGGCGGACATCTGCGAAGAGGCCAAC	1899
Qy	680	LeuSerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArg	699
Db	1896	CTGAAGCC-----ATCAACATACAGCGGCGCGACGCTGACGTCACCCCAACATC	1946
Qy	700	GlyProHis---PheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluVal	718
Db	1947	GAGCCCTTACGTTCTGAGCTGCCCTTTGTCCCGGCGCGGTGCGGAAGAACTGGACATC	2006
Qy	719	SerLysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArg	738
Db	2007	ACCGCGCTGAACGGTGACATGATGCCAACACTGACCTTGCGGATCTGTGACCTGGAGGCGGG	2066
Qy	739	AlaTyrValValLeuIleArgIleAsnAspGlyLysArgProProLeuGluGlyIleVal	758
Db	2067	ATGTATGACGTCCTCCCATCTCTGACAGACTCTGGAAACCTCCCTGTGCCAACAGCTCC	2126
Qy	759	SerLeuProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlnHis	778
Db	2127	ATCATCAAAAGTCAAGCTGTGCCCATGTGATGACAAAGGGAGTGCACCAACCATTTGCC---	2188
Qy	779	GlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuVal	798
Db	2184	-----GCAGTGGCAGCGGCTGTCTGGGCACCGGAGGCATGTGGCCATCTCATCTGC	2233
Qy	799	IleGlyIleIleLeuAlaValAla-----PheIleArgIleLysIleAspLysGlyLys	816
Db	2238	ATCTCATCTCTGCTGACATGCTGTCTGTGTTGATGTGATGAAGCGGAGAGAGAG	2297
Qy	817	Asp 817	
Db	2298	GAG 2300	
RESULT 6			
US-09-643-597-126			
; Sequence 126, Application US/09643597			
; Patent No. 6426072			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Tongtong			
; APPLICANT: Fan, Liqun			
; APPLICANT: Kalos, Michael D.			
; APPLICANT: Bangur, Chaitanya S.			
; APPLICANT: Hosken, Nancy			
; APPLICANT: Fanger, Gary R.			
; APPLICANT: Li, Samuel X.			

```

: APPLICANT: Wang, Aijun
: APPLICANT: Skelky, Yasir A.W.
: APPLICANT: Henderson, Robert A.
: APPLICANT: McNeill, Patricia D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.455C11
: CURRENT APPLICATION NUMBER: US/09/643,597
: CURRENT FILING DATE: 2000-08-21
: NUMBER OF SEQ ID NOS: 369
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 126
: LENGTH: 3552
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-643-597-126

Alignment Scores:
Pred. No.: 1,556-63 Length: 3552
Score: 671.50 Matches: 187
Percent Similarity: 48.27% Conservative: 106
Best Local Similarity: 30.81% Mismatches: 263
Query Match: 15.54% Indels: 51
DB: 4 Gaps: 22

US-10-025-380-1081 (1-832) x US-09-643-597-126 (1-3552)
QY 239 TTPYBALAPROLYSPROVALGLUWETVALGUAENSERTHASPProHISPROILELYS 258
DB 476 TGG---GCACCTATTCCTTGCTCTATGCAGAGAATTCCTTGAGCCCTTCCCATGTTT 532
QY 259 ILETHIRGINVALAAGTTPASMAASPROGLYALAGINTYSELEU----- 273
DB 533 CTTCAACAAGTGGAA--TCTGATCAGACAGCAACAACTAATAGTCTTCACTCAATAAGT 589
QY 274 -----VALASPLYEGULYSEUPEOATGPHPEPROPHSERILEASPLINGULY 290
DB 530 GGACCTGGAGTTGATTAAGAACCTTAATTGTTTATATGAAGAAGAAC--ACTGGA 646
QY 291 ASPILEYRYVALTHIRINPROLEUASPARGLINGULYASPALATYR-----VALPHE 308
DB 647 AATCATATTTGGACCTCGGCTGTGGATCGTGAAAGAAATGATGTTTATTGATTGATTGCT 706
QY 309 TYRAIALVALAIALYASPOLUTYRGILYSPROLEUSERTYPROLEULULEHISVAL 328
DB 707 TATGGGTCAACTGCAGATGGATATTTCAGACAGATCTG-----CCCTCCCACTACCCATC 760
QY 329 LYSVALIYASPIILASMAASPARNPPOPTHIRCYSPROSERPROVALTHIRVALPHEGLU 348
DB 761 AGGGTAGAGATGAATAATGACACACCACTGTTTCACAGAAAGCAATTATTAATTGAA 820
QY 349 VALINGIUAENGLUARGLEULYANSEIRILEGIYTHIRLEUTHRALHISAPARGAP 368
DB 821 GTTTTGGAAAGTAGTAGACCTGTACTACAGAGGGGGGTGTTGTGCCACAGACAGAGAT 880
QY 369 GLUGIUAENTHIRLAASNSERPHLEUENANTYTRAAGLILEVALGLUGIUTHIRPROLYSEU 388
DB 881 GAACCGGACACAAATGCATACGGCCTTGAAATACACATTTCACAGACAGACCAACAGCTCA 940
QY 389 PROMETASPGIYLEUPHELEULILEGINHTRYRALAGLMELEUENIENUALYSGIN 408
DB 941 CCT-----GGGCTCTTTTCTGTGCAATCCACAGACGGCGTATATCCACAGCTCTCAT 994
QY 409 SERLEULYSLYGLINASPETHIRPROGLINTYRAENLEUTHIRILEGLUVALSERAPLYASAP 428
DB 995 TATTTGGACAGAGAGCTGTGACAGAGTACTCATTTGATATATGAAGTACAGACATGAGAT 1055
QY 429 PHELYS-----THIRLEUQSPHEVALGLINTLEASVALILEASP 441
DB 1055 GGCCAGTTTTTGGATTGATAGGACACATCAACTGTATC-----ATAACAGTAAACAGAT 110
QY 442 ILEAENASPGINILEPROILEPHEGLIUYSSERRAPTYRGILYASNEUTHIRLEUALAGLU 461

```

```

Db      1109 TCAATGATGATGACCCACCTTTCAGACAAAATGCTTATGAA-----GCATTGTGAG 1162
Qy      462 AspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluProPhe 481
      1163 GAAATGCAATTCATGCGAAATCTTACGAAATACCTATAGAAAGTAAAGATTTAATTAAAC 1222
Qy      482 ThrGlySerSerGlyIleLeuThrIleIleLeuGlyAspSerGluGlyValGluGly 501
      1223 ACTGCCAATTCGAGAGTCAATTTTACCATTTTAAAGGAAATGAAATGACATTGCAAA 1282
Db      502 ValAspThrAspProHisThrAsnThrGlyTyrValIleIleLeuGlySerProLeuAspPhe 521
Qy      1283 ATCGGCACGACGACAAAGAACTAAAGAAAGTCTTCTGTTGTAAGCCACTGAATAT 1342
      522 GluThrAlaIleValSerAsnIleValPheGlyAlaGluAsnProGluProLeuValPhe 541
      1343 GAGGAAACCGTCAAGGAACTGGAATTCGAAATTCGAAATTCGAAAGCCATTGCTGCA 1402
Qy      542 GlyVal--LysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
      1403 GATATTCACAGAGTACAGCGCTTGACAGAGCTTGCTGACATTCATGTCAGGAGATCTG 1462
Qy      561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
      1463 GATGAGGAGCGCTGATGACACTCTGACAGCCCAATATGTCGCGATTTAAAGAACTTGA 1522
Qy      581 IleGlyThrIleValGlyAsnValThrAlaLysAspProGlu-----GlyLeuAsp 597
      1523 CTGGGGTCAAGATCAACGCGCTATAGCGCATATACCCGAAATAGAAATGGCAATGCT 1582
Qy      598 IleSerTyrSerLeuArgGlyAspThrArgGlyTyrLeuLysIleAspHisValThrGly 617
      1583 TTAAAGTACAAATAATGATGATCTTAAGGTTGATCACCATTGATGAATTTCAAGGG 1642
Qy      618 GluIlePheSerValAlaProLeuAspArgGluAlaLysPhePro-----Tyr 633
      1643 TCAATCTAACTTCCTCAAAATCTCGATATGAGAGGTTGAACTCCCAAAAATGACTGTAT 1702
Qy      634 ArgValGlnValAlaIleThrGluValGlyGlySerLeuSerSerValSerGluPhe 653
      1703 AATATTACAGTCTCCGCAATAGACAAAGATGATGATGATGATGATGATGATGATGATG 1762
Qy      654 HisLeuIleLeuMetAspValAsnAspAsnProProArgLeuAlaLysAspTyrThrGly 673
      1763 AAC-----ATTGAAGATGTAAATGATATCCACCAAAATACTTCAAGATAT----- 1810
Qy      674 LeuPhePheCysHisPheLeuSerAlaProGlySerLeuIlePheGluAlaThrAspAsp 693
      1811 GTAGTCATTTGCAAAACCAAAATGGGGTATACCGACATTTTAA-----GCTGTTGATCT 1864
Qy      694 AspGlnHisLeuPheArgGlyProHisPheThrPheSerLeuGlySer----- 711
      1865 GATGAACCTGTCATGAGAGCTCA--TTTATTTTCAGTTTGCCCAATCTTCCAGAA 1921
Db      712 LeuGlnAsnAspIleGluValSerGlyIleAsnGlyThrHisAlaArgLeuSer---Thr 730
Qy      1922 ATCAGTAGACTGTGAGAGCTCACCACAAAGTTAATGATGATGATGATGATGATGATGATG 1981
Qy      731 ArgHisThrAspPheGluGluArgAlaTyrValValLeuIleArgIleAsnAspGlyGly 750
      1982 AAAAATGCTGATTTTCAGAA-----TATACCATTCCTATTACTGTAAGAAC----- 2029
Qy      751 ArgProProLeuGluGlyIleValSerLeuProValThrPheCysSerCysValGluGly 770
      2030 AGGAGCCGCGCAAGCTGCAACAAATTTATGAGATTAATCTGTGTGATTAATCTCATCA 2089
Qy      771 SerCysPheArgProAlaGlyHisGlnThrGlyIleProThrValGlyMetAlaVal--- 789
      2090 ACTCAGTGTGTGCGATTCAGAGAGTACAGAGTATACTTGGAAATAGGCGATCTCT 2149
Qy      790 GlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAlaValPheIleArg 809
      2150 GCAATATTACTGGGTATAGCACTGCTCTTTCTGTATGCTAACTTATGATGAGAGTT 2209

```

```

Qy      810 IleLysLysAspLysGlyLys 816
Db      2210 TTGTGTCACATTAAGGAAA 2230

RESULT 7
US-08-188-228-61
; Sequence 61, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,228
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,460
; FILING DATE: 19 Apr 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 Apr 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5597725and, Greta E.
; REGISTRATION NUMBER: 35,302
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2690 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-188-228-61

Alignment Scores:
Pred. No.: 1,596-62 Length: 2690
Score: 660.50 Matches: 205
Percent Similarity: 42.31% Conservative: 103
Best Local Similarity: 28.16% Mismatches: 288
Query Match: 15.29% Indels: 132
DB: 1 Gaps: 23

US-10-025-380-1081 (1-832) x US-08-188-228-61 (1-2690)

Qy      128 GlnSerLysTyrGluGlySerValArgGlnAsnSerArgPro-GlyLysPheProPheLeu 147
      441 CAAATGACGCGCCAGAACTCCGCTCTGTCGTCCTGTCCTGTCCTGTCCTGTCCTGTC 500
Qy      147 rValAsnAlaThrAspLeuAspAspProAlaThrPro----- 159
      501 AACATCTGAGAGAAATTTGAC-----TGCACCTCTGATTTACACGAAAGTGTCCA 554
Qy      159 ----- 159

```

Db 555 TATCATCAGCAGCTGTAATTCATTGAGGACCACTGATCTTAACCTTGACCTTCAGTGA 614
 Qy 160 -----AsnGlyLeuLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAs 176
 Db 615 CTGTAAAGGAAACGACAGACTACCGCTATGAGGTC-----TC 650
 Qy 176 nValMetTyrPheGlnIleAsnAsnLeuTyrGlnValIleSerLeu-----ThrAr 193
 Db 651 GAGCCCATCTTCAAGGTGAACAGCATGGCGGCTTGTTGCTCTGAGAAACATTAATCTGC 710
 Qy 193 gGluGlySerGlnIleLeuAsnProAlaTyrAsnProSerTyrAsn-----208
 Db 711 AGTGGGCAAAACTGTGTTCCATGCATGCAGACCCCATCGGAGATATGCAAGACT 770
 Qy 209 -LeuValIleSerValIleAspMetGlyGlyGlnSerGluAsn-----SerPheSerAs 226
 Db 771 CGTATGTCGGGGGAAAGACATCCAGGCTCCTTCAGAGATATTAATTAATTTGCAAG 830
 Qy 226 pThrThrSerVal-----AspIleIleValThrGluAsnIleTyrVal 241
 Db 831 AACTTCTCTCTCCCAAGACAAAGAGGTCATTTGTGTATCT-----873
 Qy 241 aProLysProValGluMetValGluAsnSerThrAspProHisProIleuTyrGlnIle 261
 Db 874 -----CCCATTTTAATTCAGAGAAATCAGAGACAGCTTCCCAAGAGATGTTGGCAA 926
 Qy 261 nValAlaTyrPheAsnAspProGlyAlaGlnTyrSerLeuValAspGlySerGluValPro 281
 Db 927 GGTG-----GTGATGTGTCAGGCGCAGAAAG 953
 Qy 281 gPheProPhe-----SerIleAspGlnGlu-----289
 Db 954 GTCCAGATTCGGGCTCACTGAAAGGAGGTGATCAAGAGCTTAAGAGATTTTCAGAAAT 1013
 Qy 290 -----GlyAspIleTyrValThrGlnProLeuAspArgGlnGluTyrAspAl 305
 Db 1014 CAATGAGAACACAGGAGCGTCTCCGTGACAGGACCTTGACAGAGAAATGATTCGCTGT 1073
 Qy 305 aTyrValPheTyrAlaValAlaTyrAspGlnTyrGlyLysProLeuSerTyrProLeuG 325
 Db 1074 TTATCAACTATTTGTGGAGACCACTGATGTCAATGGCAAACTCTCGAGGGCGGCGGCC 1133
 Qy 325 uIleHisValIleValIleAspIleAsnAspAsnProProThrCysProSerProValTh 345
 Db 1134 TCTGAAGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
 Qy 345 rValPheGlnValGlnGluAsnGluAsnGlyAsnSerIleGlyThrIleuThrAla 365
 Db 1194 CATGGGCGACGTGATGAGAGGTGACCCACAGGACCAAGTATGGGATGACAGCTT 1253
 Qy 365 sAspArgAspGlnGluAsnThrAlaAsnSerPheLeuAsnTyrArgIleValGlnGlnTh 385
 Db 1254 TGATGAGATGACCCAGCCAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
 Qy 385 rProIleuProMetAspGlyLeuPheLeuIleGlnThrTyrAlaGly--MetLeuG 404
 Db 1314 GCTGCAAGACCATCTCCCAACATGTTTACATGATCTGAGAAAGAGACATTTGATC 1373
 Qy 404 nLeuAlaLysGlnSerLeuTyrLysGlnAspThr-----ProGlnTyrAsnLeuTh 421
 Db 1374 TGTGTGTGACCTGCGGTGAGACCGAGACCTGGAATAATCCCAAGTATGAGAT 1433
 Qy 421 rIleGlnValSerAspLys-----AspPheLysThrLeuCysPheValG 436
 Db 1434 CATGAGGCTCMAATATGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1493
 Qy 436 nIleAsnValIleAspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAs 456
 Db 1494 GATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558
 Qy 456 nLeuThrIleuAlaGluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAs 476
 Db 1549 -CAAGCCACAGTCAAG 1604

Qy 476 rAlaAspGluProPheThrGlySerSerLysIleLeuTyrHisIleIleLeuGlyAspSe 496
 Db 1605 TAAAGATGACCCACACAGATGATGAGAGGCTGCTTACATCATATCAACGAAACCC 1664
 Qy 496 rGluGlyArgLeuGlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLeu 516
 Db 1665 CGGGCAGAGCTTGAATCAACACCAACCTCAACCAACCAAGAGGATGCTTCTGTGT 1724
 Qy 516 sLysProLeuAspPheGlnThrAlaAlaValSerAsnIleValPheLeuAlaGluAsnPr 536
 Db 1725 CAACCATGAGCTATGAATGAATTTCTGCTTCCACACCTGCTATCAAAAGTGAATAATGA 1784
 Qy 536 oGluProLeuValPheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuI 556
 Db 1785 AGACCACTCTGTACCCAGACTTCTTACCGGCCCACTCCACAGCCATCCATCATC 1844
 Qy 556 eValThrAspValAsnGlnValaProGlnPheSerGlnHisValPheGlnAlaLysValSe 576
 Db 1845 TGTCTGAGATGTCAGAGAGGCGCCAGTCTTACCCAGACCCCATGATGGTGACAGGCA 1904
 Qy 576 rGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLe 596
 Db 1905 GAGGAGCTCTCTGTGGGCGAGCGTGTGTGACAGAGTATGACAGGAGCCGACTCTCT 1964
 Qy 596 uAsp-----IleSerTyrSerLeuArgGlyAspThrArgGlyTyrLeuLysIleAs 613
 Db 1965 GACGACATCAACCATCAGATGATTTCTGTTTCAAGACCCACAGAGTTGGCTGAATATTA 2024
 Qy 613 rHisValThrGlyGluIlePheSerValAlaProLeuAspArgGluAla-----629
 Db 2025 CCCCATCAATGAGACCTGTTGACACACAGCTGTGTGAGACCGTGAATTCGCA 2084
 Qy 630 -GlySerProTyrArgValGlnValAlaThrGluValGlyLysSerSerLeuSerSe 649
 Db 2085 CAACAGGTGTACACTGCTCTCTGCGCAATTTGACAGTGGCAACCTCCGCTACGGG 2144
 Qy 649 rValSerGluPheHisLeuIleLeuMetAspValIleAspAsnProProArgLeuAla 669
 Db 2145 CACTGGGACTTCTGTGATTAACCTTGAGAGCGATTAAGATCTTACCGCCG-----2193
 Qy 669 sAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaPro-----683
 Db 2194 -----TTCAATTAACCCACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2234
 Qy 684 -----GlySerLeuIlePheGluAlaThrAspAspArgGlnHisLeuPheArgGlyPr 701
 Db 2235 AAACCTCAGTGTGATCATTTTGGAGAGCATGATTAAGATCTTACCGGAAATACAGATCC 2294
 Qy 701 oHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAsp---TrpIleValSerLys 720
 Db 2295 T---TTCAATTTGAAATTCACAAACAGAGCTGTCTCTGATTAAGTCTGGAAGATCTCCA 2351
 Qy 720 sIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGlnGluAlaGlyAla 740
 Db 2352 GATCAACAAATACACAGCCCTGGTAAGCTTTTCA---AATGTGAACAAACAAACATA 2408
 Qy 740 rValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyLysIleValSerLe 760
 Db 2409 CAACCTGCCATCATGATGTGACAGATTCAGGAAACCAACCATGACGAATATCACAGATCT 2468
 Qy 760 uProValThrPheCysSerCys 767
 Db 2469 CAGGATCAAGATGCTCTCTGC 2490

RESULT 8

US-08-332-643-55

; Sequence 55, Application US/08332643

; Patent No. 563634

; GENERAL INFORMATION:

; APPLICANT: Suzuki, Shintaro

; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS

; NUMBER OF SEQUENCES: 56


```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESSEE: Bicknell
STREET: Two First National Plaza, 20 South Clark
STREET: Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,643
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/872,643
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5639634 and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30795
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 2690 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-332-643-55

Alignment Scores:
Pred. No.: 1,596-62 Length: 2690
Score: 660.50 Matches: 205
Percent Similarity: 42.31% Conservative: 103
Best Local Similarity: 28.16% Mismatches: 288
Query Match: 15.29% Indels: 132
DB: 1 Gaps: 23

US-10-025-380-1081 (1-832) x US-08-332-643-55 (1-2690)
QY 128 GlnserlyrTyrglugliserValarglnasnsrArPro-glyAsrProphleuTy 147
D 441 CAAATGCAAGCCGAGAACTCCGCTGTTCTGCTGCTTCTCTGCTCCAGATGCTGCT 500
QY 147 rValasnaIaThraspleuAspProalaThrPro- 159
D 501 AACATGTCAGAGATTGGAC-----TGCACCTCGATTGACAGATTGACAGAAAGTTTCA 554
QY 159 ----- 159
D 555 TATCATCAAGCCAGCAAGTAATTCATTGAGACCAAGTCAATTCATACTTGAAGCTTCAAGTGA 614
QY 160 -----AsnIyGlnleuTyTyrglnleValIleGlnleuPrometIleasna 176
D 615 CTGTAAGGAGAAAGACAGCTACGCTATGAGTC-----TC 650
QY 176 nValMeTyTyPheGlnIleasnaenlyrThrglyAlaIleSerleu-----Thrar 193
D 651 GAAGCCATCTTCAGAGTGAACAGCGATGCGGCTTGCTTGAGAAACATACTGCTGC 710
QY 193 ggluglySerGlnIleuAsnProalaTyAsnProSerTyAsn----- 208
D 711 AGTGGGCAAAACTCTGTTGCTCATGACGAGACCCCATGCGGAGATATGCGAGACT 770
QY 209 -leuValIleSerVallyAspMetGlyGlnSerGluAsn-----SerPheSerAs 226

```

```

D 771 CGTATGTCGGGGGAGAAAGACATCCAGGGCTCTTGACAGATATTAATTAATTTGCAAG 830
QY 226 pThrThsSerVal-----AspIleIleValThrglnsnleThrpySaI 241
D 831 AACTCTCTCTGTCACAGACAAAGAGTCATGTTGATVCT----- 873
QY 241 aProlySProValGluMetValGlnAsnSerThrasPProhIAspProIleThrgI 261
D 874 -----CCCATTTTATTCAGAGAAATCAGAGACCCCTTTCCCAAGATATGTCGA 926
QY 261 nValArgTyPAsnaSpproGlyValaglnTySerleuValAspIlySgluAsnProAr 281
D 927 GGRA-----GTCGATAGTACAGCGCCAGAAAG 953
QY 281 gPheProPhe-----SerIleAspGlnIu----- 289
D 954 GTCGAAGTTCGGCTCACTGGAAGAGAGATGATCAAGACCTTAAGAAATTTTCAGAAAT 1013
QY 290 -----GlyAspIleTyValThrglnProleuAspArgIglulysAspAl 305
D 1014 CAATGAGAACACAGGAGAGCTCTCCGTGACAGGACCTTGACAGAGATATTCGCTGT 1073
QY 305 aTyValPheTyValaValaIalysAspGlyTyrglyTySProleuSerTyPProleuG 325
D 1074 TTAATCAACTATTTGTGGAGACCACTGATGTCATGGAATAAATCTCGAGGGCGCGGTGCC 1133
QY 325 uIleHlSVallyVallyAspIleAsnaSnaPProPProThCySProSerProValTh 345
D 1134 TCTGGAAGTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
QY 345 rValPheGlnIleValGlnIleAsnIleuTyrglyTyAsnSerIleGlyThreuthrAlaH 365
D 1194 CATCGCCACGTATATGAGAGAGGTCACCCACAGGACACACAGTATGCGATGACAGCTT 1253
QY 365 sAspArGAspGlnIleAsnThrlaAsnSerPheleuAsnTyTyrgIleValIglulnTh 385
D 1254 TGATGCAAGATGACCCAGACCAAGATATGCTGCGGTATATATATATATATATATATATAT 1313
QY 385 rProlySleuPProMetAspIlyleuPheleuIleGlnIleThrTyAlglIy---MetleuG 404
D 1314 GCCTGACAAAGCCATCTCCCAACATGTTCTAATCATGATCTCTGAGAAAGAGACATTTGCTAC 1373
QY 404 nleuAlaIyGlnSerleuIlySgluAspThr-----ProGlnIlyrAsnleuTh 421
D 1374 TGTGTGTCACTCGCTGCTGAGACGAGAGACTCTGGAATAATCCAAATATGAACTGAT 1433
QY 421 rIleGlnIleSerAspIlyS-----AspPheIlyThreucySpheValG 436
D 1434 CATCGAGGCTCAAGATATATGCTGACATGATGATGATTAACGACGACGACGACGACGAC 1493
QY 436 nIleAsnValIleAspIleAsnaSpglnIleProllePheGlnIlySerAspTyrglyAs 456
D 1494 GATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
QY 456 nleuThrlleuAlaGluAspThrasnIleGlySerThrlleuThrlleGlnIlaThras 476
D 1549 -CAAGCCACAGTCGAGAGAGAGAGAGCTGTGGA---GTTATGTCATATTTGACAGTTGAAG 1604
QY 476 palAspGlnPProPheThrglySerSerIlyleuTyTyrgIleIleIleIlyAspS 496
D 1605 TAAAGATGACCCACACACAGAGTCATGAGAGGCTGCTTACATCATATCAACGAGAAACC 1664
QY 496 rgluglyArGleuGlyValAspThrasPProhIAsnThrgIlyTyValIleIleIly 516
D 1665 CGGGCAGAGCTTGAATATCACAACCACTCAACCAAGAGGATCTTCTGTTGT 1724
QY 516 slySProleuAspPheGlnThrlaIalAserAnIleValPheTyAlaGluAsnPr 536
D 1725 CAACCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
QY 536 ogIuPProleuValPheGlyVallyTyAsnaIAserSerPheAlaIyPheThreul 556

```

Db 1785 AGACCCAGCTGTAACCGACCTCTCTAAGCGCCAGCTCCAGACCAACCTCCACATCAC 1844
 Qy 556 eValThAspValAsnGlnAlaProGlnPheSerGlnHisValPheGlnAlaValSer 576
 Db 1845 TGTCCTGGAAGTCAACGAGGCGCCAGTCTTCTTACCCAGACCCCATGATGTGACAGGCA 1904
 Qy 576 rGluAspValAlaIleGlyThrIysValGlyAsnValThrAlaIysAspProGluGly 596
 Db 1905 GGAAGACCTCTCTGTGGGACGCGCTGCTGACAGTGAATGACAGGACCCCGACCTCCT 1964
 Qy 596 uAsp-----IleSerTySerLeuArgGlyAspThrArgGlyTyrLeuIysIleAs 613
 Db 1965 GCAGCATCAAAACCATGATCTCTGTTTCAAGAGACCCAGAGGTGGCTGATATATAA 2024
 Qy 613 pHisValThrGlyGluIlePheSerValAlaProLeuAspArgGluAla----- 629
 Db 2025 CCCCATCAATGGGACTGTGTGACACACAGCTGCTGACCGTGAAGTCCCATTTGTGCA 2084
 Qy 630 -GlySerProTyrArgValGlnValAlaThrGlnValGlyGlySerSerLeuSerse 649
 Db 2085 CAACAGCGGTGACCTGCTCTCTCTGCGCAATTGACAGTGGCAACCTCCCGCTAGCGG 2144
 Qy 649 rValSerGluPheHisLeuIleLeuMetAspValAsnAspAspProArgLeuAla 669
 Db 2145 CACTGGAGCTTGTCTGATTAACCTCGAGAGAGTGAATGACATGCCCG----- 2193
 Qy 669 sAspTyrThrGlyLeuPhePheCysHisProLeuSerAlaPro----- 683
 Db 2194 -----TTCAATTACCCCAAGTGAAGTGAAGTGTGATGATGCCAA 2234
 Qy 684 -----GlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGlyPr 701
 Db 2235 AAACCTGAGTGTACTATTGTTGGAGCATCGATGAGATCTTCAACCGAATACAGATCC 2294
 Qy 701 oHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAsp--TyrGluValSerly 720
 Db 2295 T--TTCAATTGAAATCCCAACAAAGCTGTCTCGATTAAGTCTGGAAGATCTCCAA 2351
 Qy 720 gIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGlnAlaGly 740
 Db 2352 GATCAACAATGACACGCCCTGTGACCTTTCTTCAA--AATGTGAACAAAGCAACTA 2408
 Qy 740 rValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluIleValSerle 760
 Db 2409 CAACCTCCCATCATGTGTGACAGTTCAAGGAACCAACCATGACGAATTCACAGATCT 2468
 Qy 760 uProValThrPheCysSerCys 767
 Db 2469 CAGGGTACAAGTGTCTCTGC 2490
 RESULT 9
 US-08-332-638-61
 ; Sequence 61, Application US/08332638
 ; Patent No. 5646250
 ; GENERAL INFORMATION:
 ; APPLICANT: Suzuki, Shintaro
 ; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Borun
 ; STREET: 6300 Sears Tower, 233 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/332,638

; FILING DATE: 01-NOV-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/872,643
 ; FILING DATE: 17 APR 1992
 ; APPLICATION NUMBER: US/08/049,460
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5646250and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (312) 474-6300
 ; TELEFAX: (312) 474-0448
 ; TELETYPE: 25-3856
 ; INFORMATION FOR SEQ. ID NO: 61:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2690 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; US-08-332-638-61
 Alignment Scores:
 Pred. No.: 1,596-62 Length: 2690
 Score: 660.50 Matches: 205
 Percent Similarity: 42.31% Conservative: 103
 Best Local Similarity: 28.16% Mismatches: 288
 Query Match: 15.29% Indels: 132
 DB: 1 Gaps: 23
 US-10-025-380-1081 (1-832) x US-08-332-638-61 (1-2690)
 Qy 128 GlnSerLysTyrGluGlySerValArgGlnAsnSerArgPro-GlyLysProPheLeu 147
 Db 441 CAATAAGCAAGCCAGCAAGTCCGCTCTGTGCGCTTCTCTGCTCCAGGTGCTGCT 500
 Qy 147 rValAsnAlaThrAspLeuAspAspProAlaThrPro----- 159
 Db 501 AACATCTGCAAGAAATTGGAC-----TGCATCTCTGATTTACAGCAAAAGTGTCCA 554
 Qy 159 ----- 159
 Db 555 TATCATACAGCCAGCGAATTGATGAGACAGTCAATTCTAACTTGACCTTGATGA 614
 Qy 160 -----AsnGlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAs 176
 Db 615 CTGTAAAGGAAACGACAGCTACGCTATGAGTGC-----TC 650
 Qy 176 nValMetTyrPheGlnIleAsnAsnLysThrGlyAlaIleSerLeu-----ThrAr 193
 Db 651 GAGCCATTAATCTTAAGGTGAACAGCATGGCGGCTTAAGTGTCTGAGAAACATAACTGC 710
 Qy 193 gGluGlySerGlnGluLeuAsnProAlaIysAsnProSerTyrAsn----- 208
 Db 711 AGTGGGCAAAACTCTGTGTCATGACGAGACCCCATGGGAGATATGCGAGAACT 770
 Qy 209 -LeuValIleSerValIysAspPheClyGlyGlnSerGluAsn-----SerPheSerAs 226
 Db 771 CGTGATTTGCGGGGGGAAAGACATCCAGGCTCTCTTGAGGAATATTAAATTTGCAAG 830
 Qy 226 pThrThrSerVal-----AspIleIleValThrGluAsnIleTyrVal 241
 Db 831 AACTTCTCTGCTCCCAAGCAAAAGAGTCCATTTGTGTAATCT----- 873
 Qy 241 aProLysProValGluMetValGluAsnSerThrAspProHisProIleValIleThrG 261
 Db 874 -----CCCATTTTAATTCAGAGATCAGAGACAGCGCTTTCCAAAGATGTGGCAA 926
 Qy 261 nValArgTyrPheAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuProAr 281
 Db 927 GGTAA-----GTCGATGATGACAGCGCCAGAAAG 953

QY 281 gphepPrope-----SerileAspGlu-----289
 Db 954 GTCCAGTTCGGCTCACTGAAAGAGATGATCAAGAGCTTAAGAAATTTTTCAGAT 1013
 QY 290 -----GlyAspIleTyValThrGlnProLeuAspArgGluGlyAspAl 305
 Db 1014 CAATGAGACACAGAGAGCGCTCCGTGACACGAGCTTTGGACGAGAAAGTAAATCGCTGT 1073
 QY 305 aTyValPheTyValAlaValAlaValAspGluTyGlyValProLeuSerTyProLeuG 325
 Db 1074 TTATCAACTATTTTGGAGACCACTGATGTCAATGGCAAACTTCCAGGGGCGGTGCC 1133
 QY 325 uileAsValIleValIleAspIleAspAspProProCysProSerProValTh 345
 Db 1134 TCTGAGATCATTTGTGATTGATCAAAATGACACGACCGATCTTTGGGGAAGGCCCTTA 1193
 QY 345 rValPheGluValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 365
 Db 1194 CATGGCCACGTCATGAGAGGTCACCCACAGCCACACAGTATGCGGATGACAGCTT 1253
 QY 365 sAspArgAspGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 385
 Db 1254 TGATGCAATGATGACCCACCCACCATATGCTTCTCGGTATATATCGGTCAACAGAC 1313
 QY 385 rProTyIleuProMetAspGlyIleuPheLeuIleGlnThrTyValGly---MetLeuG 404
 Db 1314 GCGTCAGAACCCATCTCCCAACATGTTCTATCATGATCTGAGAAAGAGACATGTGCAC 1373
 QY 404 nleuAlaValGlnSerIleuTyIleuAspThr-----ProGlnTyIleuLeuTh 421
 Db 1374 TGTGTGTCACTGCGCTGAGACGACAGACTCTGAAATCCCAAGTATGAAATCGAT 1433
 QY 421 rIleGluValSerAspIle-----AspPheIleThrIleuCysPheValG 436
 Db 1434 CATGAGGCTCAAGATATGCTGAGCTGATGTTGATTAACAGGACCGGACACAGCCAC 1493
 QY 436 nileAsnValIleAspIleAsnAspGlnIleProIlePheGlnIleuSerAspTyGlyAs 456
 Db 1494 GATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
 QY 456 nleuThrIleuAlaGlnAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAs 476
 Db 1549 -CAAGCCACAGTCGAGGAGAGAGCTGTGGA---GTATTTGATCAATTTGACAGTGAAGA 1604
 QY 476 rAlaAspGluProPheThrGlySerSerIleLeuTyIleIleIleGlyGlyAspSe 496
 Db 1605 TAAAGATGACCCCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1664
 QY 496 rGlnGlyArgLeuGlyValAspThrAspProIleThrAsnThrGlyTyValIleIleIle 516
 Db 1665 CGGGCAGAGCTTGAATCAACACCAACCTCAACCAACCAAGAGATGCTTTCTGTGT 1724
 QY 516 slyAspProLeuAspPheGlnThrAlaAlaValSerAsnIleValPheIleValGlnAsp 536
 Db 1725 CAATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
 QY 536 ogIuProLeuValPheGlyValIleTyIleAsnAlaSerSerPheAlaIlePheThrIleu 556
 Db 1785 AGACCCACCTGTCACCGACCTCTCTACCGGCGGCTTCCACAGCCACCGTCCACATCAC 1844
 QY 556 eValThrAspValAsnGlnAlaProGlnPheSerGlnIleValPheGlnAlaValSer 576
 Db 1845 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 QY 576 rGlnAspValAlaIleGlyThrIleValGlnValGlnValGlnValGlnValGlnValGln 596
 Db 1905 GAGAGACCTCTCTGCTGAGAGCTGTGACAGATGATGATGATGATGATGATGATGATGAT 1964
 QY 596 uAsp-----IleSerTyIleuAspGlyAspThrArgGlyThrIleuTyIleAs 613
 Db 1965 GCAGCATCAAAACCATGAGGATATCTGTTTCAAGAGACCCAGAGGTTGGCTGAATATTA 2024

QY 613 pHsValThrGlyGlnIlePheSerValAlaProLeuAspArgGluAla-----629
 Db 2025 CCCCATCAATGGAGCTGTGACACACAGCTGTGTCGACCGTGAAGTCCCCATTTTGA 2084
 QY 630 -GlySerProTyArgValGlnValAlaIleThrGlnValGlyGlySerSerIleuSerSe 649
 Db 2085 CAACAGGTGTACATGCTCTCTTCCGCAATGTACAGTGGCAACCTCCGCTACGGG 2144
 QY 649 ValSerGluPheIleIleuIleuMetAspValAspAspProProArgIleuAlaI 669
 Db 2145 CACTGGACCTTGTGTATACCTTGGAGACGATGATGATGATGATGATGATGATGATGATGAT 2193
 QY 669 sAspTyThrGlyIleuPhePheCysHisProLeuSerAlaPro-----683
 Db 2194 -----TTCAATTAACCCACAGTATGAGATGCTGTGATGATGCCAA 2234
 QY 684 -----GlySerIleuIlePheGlnAlaThrAspAspAspGlnHisIleuPheArgGlyPr 701
 Db 2235 AAACCTCAGTGTGATTTTGGAGCATGATGATGATGATGATGATGATGATGATGATGAT 2294
 QY 701 ohIlePheThrPheSerIleuGlySerGlySerLeuGlnAspAsp---TrpGluValSerTy 720
 Db 2295 T---TTCAATTTGAAATTCACAAACAGCTGTTCCTGATTAAGTCTGGAAGATCTCAA 2351
 QY 720 sIleAsnGlyThrHisAlaArgIleuSerThrArgHisThrAspPheGlnGluAlaTy 740
 Db 2352 GATCAACAATACACAGCCCTGTGATGCTTCA---AACTGAACAACCAACTA 2408
 QY 740 rValIleuIleArgIleAsnAspGlyIleArgProProLeuGlnGlyIleValSerIle 760
 Db 2409 CAACCTGCCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
 QY 760 uProValThrPheCysSerCys 767
 Db 2469 CAGGTACAAATGCTGCTCTGC 2490

RESULT 10
 US-08-474-067-3
 Sequence 3, Application US/08474067
 Patent No. 5811518
 GENERAL INFORMATION:
 APPLICANT: Ranscht, Barbara
 TITLE OF INVENTION: T-Cadherin Adhesion Molecule
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell and Flores
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,067
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/213,361
 FILING DATE: 14-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/607,293
 FILING DATE: 30-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1682
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949

```

: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 2779 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
: FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 40..2191
:   US-08-474-067-3
:
: Alignment Scores:
:   Pred. No.: 3,44e-59
:   Score: 630.50
:   Percent Similarity: 44.77%
:   Best Local Similarity: 28.05%
:   Query Match: 14.59%
:   Gaps: 25
:
US-10-025-380-1081 (1-832) x US-08-474-067-3 (1-2779)

```

```

QY 150 AlaThrAspLeuAspProAlaThrProAsnGlyGln-----LeuTyrTyrGlnIle 167
DB 103 GCAGAAAGATTAGAA-----TGACCCCTGATTCAGCAAAAGTTTTTAT----- 150
QY 168 ValIleGlnLeuProMetIleAsnAsnValMetTyrPheGlnIleAsnAsnLeuThrGly 187
DB 151 ---ATTGAACAGCCA-----TTGAAATTCACAGACGACCA 186
QY 188 AlaIleSerLeuThrArgGlu-----GlySerGlnIleAsn---ProAlaIys 203
DB 187 ATTGTGAACCTGCTGTTGATGACTGCAAGGGAAATACAAATTGAATCTGAAGTTCT 246
QY 204 AsnProSerTyrAsnLeu-----ValIleSerValIysAspMetGly--- 217
DB 247 AACCCAGACTTAAAGTGAACAGATGATCTTAACTTGAATGCAAGAAATGATCAGA 306
QY 218 -----GlyIleSerGlnIleAsnSerPheSerAspThrThrSer 229
DB 307 GCTGGCAGAGCTTTGTTGTCATGACGAGTCTGAGCATGCT---GAGGATATGACAGA 363
QY 230 ValAspIleIleValThrGluAsnIleTyrPylAspProLys----- 243
DB 364 ATTTGATTTGTTGAGCTGATGAGAGACAGATGATTAAGAATCTTTAAGATGAA 423
QY 244 -----ProValGluMetVal 248
DB 424 GGCACCTTGAAATTCACAGACAAAGGCTATTCTGCGACCTCCAAATTTAATCCA 483
QY 249 GluIleSerThrAspProIleProIleValIleThrGlnValArgTyrAsnAsp----- 266
DB 484 GAAATACAAAGACCAACCATTTCCAGATCAGTTGGCAAGGTCAATCAGAGGAAAGGACA 543
QY 267 ProGlyIleGlnTyrSerLeuValAspLys-----GluLysLeuProArgPheProPhe 284
DB 544 GAGGAGCAAGAGTTCCGACTCTCTGCTAAGGAGAGATCAAGACCGAAAGAAATTTT 603
QY 285 SerIleAspGln---GluGlyAspIleTyrValThrGlnProLeuAspArgGluGluLys 303
DB 604 AGATCAATGAGATCAGTGGGAGATGCTCTGTGACCCGACCCCTGATAGGAAGACATA 663
QY 304 AspAlaIleTyrValPheTyrAlaValAlaIleAspGlnTyrGlyLysProLeuSerTyrPro 323
DB 664 GCCAATATATGAGCTGGAAAGTTAAAGATTTAACTGGGAAATATCATGATGATGCGCCA 723
QY 324 LeuGluIleIleValIysValIysAspIleAsnAspAsnProThrCysProSerPro 343
DB 724 GTCCGCCCTAGATATTTCTGTTATGTAATTAAGATTAACAGGCGAGTTTCAAGAGAGA 783
QY 344 ValThrValIlePheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThr 363
DB 784 CCTATGTTGGTCACTGATGAAAGATCCCTACAGAGAACACTGATGATGCGATGACA 843

```

```

QY 364 AlaIleAspArgAspGluGlnIleValIleAsnSerPheLeuAsnTyrArgIleValGlu 383
DB 844 GCATTTGATGCTGATGATGATCTGACACAGACCAAGCTCTTCTGGGTATTAACATCTCAAG 903
QY 384 GlnThrProLysLeuProMetAspGlyLeuPheLeuIleGln----- 397
DB 904 CAGACACCTTACCAAACTTCCCAAAATATGTTTACATTTGACCCAGACAAAGAGATATT 963
QY 398 ---ThrTyrAlaGlyMetLeuGlnIleValAlaIleGlnSerIleValIleAspThrPro 416
DB 964 GTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
QY 417 GlnTyrAsnLeuThrIleGlnValSerAspLys-----AspPheLysThr 431
DB 1015 AAGTACAGAGTGTGTTTAAAGCAAGATATGCGCGCATGATGAGGACTTACTGGA 1074
QY 432 LeuCyAspPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
DB 1075 ACTGCAACTGCGCACTATTCTTATGATGACAAAGACACCAACCCAGAAATTTACCAAG 1134
QY 452 SerAspTyrGlyAsnLeuThrIleGlnValSerAspThrAsnIleGlySerThrIleLeuThr 471
DB 1135 AAGAGATT---CAGGCCACAGTAAAGAA---GAGTCAAGAGATTAATGTAAC 1185
QY 472 IleGlnAlaThrAspAlaAspGluProPheThrGlySerSerLysIleLeuTyrHisIle 491
DB 1186 TTAACTGCTGCTGATCCGAGATGATCCAGCAACTGGAGCATGAGAGTGTCTACATATT 1245
QY 492 IleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThrGly 511
DB 1246 ATTACGGAATATCAGGCGCAGAGTGTGAATATCAATCAATCCAGACTTAATGAGGA 1305
QY 512 TyrValIleIleLysLysProLeuAspPheGluThrAlaIleValSerAsnIleValPhe 531
DB 1306 ATGCTCTGCTGCTCAACCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 1365
QY 532 LysAlaGluAsnProGluProLeuValPheGlyValIleLysTyrAsnIleAspSerPheAla 551
DB 1366 AAAGTAAAGATTAAGAACCCGTTGATTCAGACATGAGCCATGAGCTCCAGTTCCACAGCA 1425
QY 552 LysPheThrLeuIleValIleThrAspValAsnGluAlaProGlnPheSerGlnIleValPhe 571
DB 1426 ACAGTTCAGATCAACGTTGAGATGTAAGAAAGCCCTGTTTCCACCCAAACCCAGT 1485
QY 572 GlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlnValIleThrAlaLys 591
DB 1486 ACAGTGCACAAACAGAGAAATCCCTATTGGAGCATTTGTGTTAACAGTAATGCCACT 1545
QY 592 AspProGluGlyLeuAsp-----IleSerTyrSerLeuArgIleAspThrArgGly 608
DB 1546 GATCCAGATCTTTGCAACATCAGACATCAGGATTCAGATTAAACAAGATCCAGCAAGC 1605
QY 609 TyrLeuLysIleAspHisValIleThrGlyLysIlePheSerValAlaProLeuAspArgGlu 628
DB 1606 TGCGTAGAGATTAATCCCAACCAATGATGACCGTTGCCACCACTGCTGCTGATGCGGAA 1665
QY 629 Ala-----GlySerProTyrArgValGlnValAlaIleThrGluValGlyGly 644
DB 1666 TCTCCGATGTCAGATTAACAATACATGCTCTCTCTCTGCGATATGACAGTGGTAAAC 1725
QY 645 SerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValIleAspAspPro 664
DB 1726 CTTCTGCTGATCAGATCAGAACTTTAACAATCATCATCCTTGAAGAGACGTCATTAATGCT 1785
QY 681 SerAlaProLysSerLeuIlePheGlnAlaIleThrAspAspArgGlnIleLysLeuPheArgGly 700
DB 1828 AAAGATTCAGAGTATGAGTTTGAAGACATCCAAACCACTCCATCCCAACACAGAT 1887
QY 701 ProHisPheThrPheSerLeuGlySerGlySer---LeuGlnAsnAspTyrGluValSer 719

```



```

Db      1135 AAGAGATT---CAGGCCACAGTAAAGAA-----GGAGTCAAGAGATTAATGTAAC 1185
Qy      472 TleGlnAlaThrAspAlaAspGluProPheThrGlySerSerIleLeuTYrHisIle 491
Db      1186 TTAACTTGTGGTCCAGAGATGAGCCAGCAACTGAGACATGAGAGCTGCTTACACTATT 1245
Qy      492 TleGlyGlyAspSerGluGlyArgLeuGlyValAlaAspThrAspProHisThrAspThrGly 511
Db      1246 ATTAAACGAATCCAGAGGACAGAGTTTGGAAATTCATCAATCCCGAGCTTAAGAGGA 1305
Qy      512 TyrValIleIleLeuGlyGlySerGluAspPheGluThrAlaAlaValaSerAsnIleValaPhe 531
Db      1306 ATGCTCTCTGTGTGCAAACTTGTAGACTATGAGATTTCAGAGATTTCACACTTCTGATA 1365
Qy      532 LysAlaGluAsnProGluProLeuValaPheGlyValIleGlyTyrAsnAlaSerSerPheAla 551
Db      1366 AAAGTAGAAATAGAAAGACCCGTTGATTCCAGACATAGCTGCTCCAGTTCCACAGCA 1425
Qy      552 LysPheThrLeuIleValaThrAspValaAsnGluAlaProGlnPheSerGlnHisValaPhe 571
Db      1426 ACAGTTCAGATCACCGTTGAGAGATGTGATGAGAGCCCTGTTTCCACCCAAACCCATG 1485
Qy      572 GlnAlaLysValaSerGluAspValaAlaIleGlyThrLysValaGlyAsnValaThrAlaLys 591
Db      1486 ACAGTGAACAAAGAGAGAACATCCCTATTGGCAGCATTTGTTAAACAGTAATAGCCACT 1545
Qy      592 AspProGluGlyLeuAsp-----IleSerTyrSerLeuArgGlyAspThrArgGly 608
Db      1546 GATCCAGATCTTCTTCAACATCAGACATCAGATTCAGATTTCAGTTTCAAGAGATCCAGAAC 1605
Qy      609 TrpLeuLysIleAspHisValaThrGlyGluIlePheSerValaAlaProLeuAspArgGlu 628
Db      1606 TGGCTAGAGATTAATCCCAACATGATGATCCCTGACCACTGCTGCTGGATCGGGA 1665
Qy      629 Ala-----GlySerProTyrArgValaGlnValaAlaThrGluValaGly 644
Db      1666 TCTCCGATGTTCCAGATTAACAAATACACTGCTCTCTCCGCAATAGACAGAGTAAAC 1725
Qy      645 SerSerLeuSerSerValaSerGluPheIleLeuIleLeuMetAspValaAspAspPro 664
Db      1726 CTTCTCTCTACAGTACAGAACTTTCACATCACTTGGAGAGACGTCATGACAAATGTC 1785
Qy      665 -----ProAlaGluValaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db      1786 CCTCCCTTACCCAAACATGCGCAAA-----GCTGTGATGATGCT 1827
Qy      681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspGlnHisLeuPheArgGly 700
Db      1828 AAAGATCTCAGAGATGAGTCTTAGAGCATCAGACAAAGCTCCATCCCAACACAGAT 1887
Qy      701 ProHisPheThrPheSerLeuGlySerIleSer---LeuGlnAsnAspTrpGluValaSer 719
Db      1888 CCA---TTTAAATTTGAATGAGTAAAGCAATCTGTCGCAAAAGTTATGAGATCAAC 1944
Qy      720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluAlaArgAla 739
Db      1945 AAGCTTAACAATCTCATGCTGAGAGTGTCTGCTTCA---AACCTGAAGAAAGGCCAAT 2001
Qy      740 TyrValValaLeuIleArgIleAsnAspGlyGlyArgProLeuGluGlyIleValaSer 759
Db      2002 TACAACATCCCAATCTCAGTACAGATTTCTGAAACACACTCTGATCAACACACAGAA 2061
Qy      760 LeuProValaThrPheCysSerCys 767
Db      2062 CTGAATTAACAAGTGTCTCTGC 2085

```

RESULT 12
 US-08-472-481-3
 Sequence 3, Application US/08472481
 Patent No. 5863804
 GENERAL INFORMATION:
 APPLICANT: Ranscht, Barbara

```

TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSER: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,481
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-1J 1686
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2779 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 40..2191
US-08-472-481-3
Alignment Scores:
Pred. No.: 3,44e-59 Length: 2779
Score: 630.50 Matches: 193
Percent Similarity: 44.77% Conservative: 115
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 14.59% Indels: 97
Gaps: 25
US-10-025-380-1081 (1-832) x US-08-472-481-3 (1-2779).
Qy      150 AlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTYrTYrGlnIle 167
Db      103 GCAGAAAGATTAGAA-----TCACCCCTCGAATTCACAGAAAGGTTTATT----- 150
Qy      168 ValIleGlnLeuProMetIleAsnValMetTyrPheGlnIleAsnValMetTyrGly 187
Db      151 ---ATTGAACGCA-----TTTCAATTCACAGAGACACGCA 186
Qy      188 AlaIleSerLeuThrArgIle-----GlySerGlnIleLeuAsn---ProAlaLys 203
Db      187 ATTCTGAACCTGTTGTTGATGACTGCAAGGGAATTAACAATTGAATTCGAAGTTTCT 246
Qy      204 AsnProSerTyrAsnLeu-----ValIleSerValIysAspMetGly--- 217
Db      247 AACCCAGACTTAAAGTGAACAGATGATCTTTGTTGCTGACGATGATGATGAGAA 306
Qy      218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
Db      307 GCTGGCAGACCTTTGTTGTTGCTGATGACAGCGTCTGAGCATCT---GAGGATATGACAGAA 363
Qy      230 ValAspIleIleValaThrGluAsnIleTrpIleAlaProLys----- 243

```



```

Db      364 ATTTGATGTTGGAGCTGATGAGAACGACGATCAATTAAGAAATCTTTAAGATGAA 423
      244 -----ProValGluMetVal 248
Db      424 GGCACCTTGGAATTCACAGACAAAAAGGCTATTCTGGGACTCCAAATTTAATTCGA 483
      249 GluAsnSerThrAspProHisProIleValIleThrGlnValArgTrpAsnAsp----- 266
      484 GAAATCAAGACACCATTTCCAGATGCTGGAGTTCATCAGAGTGAAGGACGA 543
      267 ProGlyValGlnIleThrSerLeuValAspIle-----GluIleuProAlaGlyProPhe 284
      544 GAGGAGCAAAAGTCCACCTCTCTGTAGGAGTATCAAGACCCGAAAGAAATTTT 603
      285 SerIleAspGln-----GluIleAspIleValIleThrGlnProLeuAspArgGluIle 303
      604 AGAATCATATGATGATGAGGAGATGCTGTGACCCGACCCCTGTGATGAGAAACATA 663
      304 AspAlaIleValPheValAlaValAlaValAspGluIleValIleValIleValIle 323
      664 GCAATATTAGAGCTGGAAGTTGAAGTAAACGATTTAAGTGGAAATCATTTGATGCCCA 723
      324 LeuGluIleHisValIleValIleValIleValIleValIleValIleValIleVal 343
      724 GTCCGCTAGATATTTCTGTATTGATCAAAATGATTAACAGCCGATGTTCAAAAGAGA 783
      344 ValIleValPheGluValGlnIleValIleValIleValIleValIleValIleVal 363
      784 CCTATGTTGTGTCACGTATGAGAAAGATCCCTTACAGAACATGTGATGCGGATACA 843
      364 AlaHisAspArgSerGluGluAsnThrAlaAsnSerPheLeuAsnIleValIleVal 383
      844 GCATTTGATGCTGATGATCTTCTAGACACAGACGCTTCTGTGCGTATTAATCTCAAG 903
      384 GlnThrProIleuProMetAspGlyLeuPheIleIleGln----- 397
      904 CAGACACCTACCAAACTCCCAAAATATGTTCTTACATGACCCAGAAAGGAGATATT 963
      398 ---ThrValIleGluMetLeuGlnIleValIleValIleValIleValIleValIle 416
      964 GTCCAGTGTGTGCTGCTGATGCTGATGCTGATGCAATG-----GAAACCCCG 1014
      417 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleVal 431
      1015 AAGTACAGCTGCTGTTATTTGAAGCCAAAGATATGGGCGCATATGCGGACCTTACGGA 1074
      432 LeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGlu 451
      1075 ACTGCAACTGCCACTATTTCTTATGATGACAAAACGACACCCAGACGATTTACCAAG 1134
      452 SerAspIleValIleValIleValIleValIleValIleValIleValIleValIle 471
      1135 AAGGAGT---CAGGCCACAGTAAAGAA-----GAACTCCAGAGTAAATTAATAC 1185
      472 IleGlnIleValIleValIleValIleValIleValIleValIleValIleValIle 491
      1186 TTAAGTCTTGTGATGACGATGACCCGACATGAGACATGAGACGCTGTCTACATATT 1245
      492 IleValIleValIleValIleValIleValIleValIleValIleValIleValIle 511
      1246 ATTAACGGAATCCAGGACAGATTTGAATATCATATCAATCCCACTCAATTAAGAGGA 1305
      512 TyrValIleIleIleValIleValIleValIleValIleValIleValIleValIle 531
      1306 ATGCTCTCTGTGCAAACTTTAGATTAAGATTTACACATTTTCAACATTTGCTGATA 1365
      532 LysAlaGluAsnProGluProLeuValPheGlyValIleValIleValIleValIle 551
      1366 AAGGTAGAAAATGAAGCCGCTGATTCAGACATACCTTACCGTCCCAAGTTCCACAGCA 1425
      552 LysPheThrLeuIleValIleValIleValIleValIleValIleValIleValIle 571

```

```

Db      1426 ACAGTTCAGATCACCGTTGAGATGTAAGTGAAGGCCCTGTTTCCACCCAAACCAANG 1485
      572 GlnAlaIleValSerGluAspValAlaIleGlyThrIleValGluValIleValIle 591
      1486 ACAGTACAAAACAAAGAAACATCCCTATTGGAGACATTTGTTAACAGTAATGCCACT 1545
      592 AspProGluIleuAsp-----IleSerTyrSerIleuArgGlyAspThrArgGly 608
      1546 GATCCAGATCTTGTGCAACATCAGATCAGATCAGGTATTCAGTTTCAAGATCCAGCAAGC 1605
      609 TrpLeuValIleAspHisValIleThrGlyGluIlePheSerValAlaProLeuAspArg 628
      1606 TGGCTAGAGTTATATCCCAACAAATGATCCGTTGCCACCATCTGCTCTGATCGGGA 1665
      629 Ala-----GlySerProTyrArgValGlnValAlaIleThrGluValIleGly 644
      1666 TCTCCGATGTTGAGATTAACAATACACTGCTCTCTTCTGGCAATAGACAGTGTAC 1725
      645 SerSerLeuSerSerValSerGluPheHisIleuIleuMetAspValAsnAspAsnPro 664
      1726 CTTCTGTACAGTACAGTACAGGAATTTACATCATCATCTGGAGAGACGTCATGACATGTC 1785
      665 -----ProArgLeuAlaIleAspTyrThrGlyLeuPhePheCysHisProLeu 680
      1786 CCTCCCTTACCCCAACACTGGCMAA-----GTCGTGTATGATGCT 1827
      681 SerAlaProGlySerIleuIlePheGluAlaIleThrAspAspGlnHisIleuPheArgGly 700
      1828 AAGATCTCAGAGTATGCTTCTTACGAGCATCAGACAAAGACCTCCATCCCAACACAGAT 1887
      701 ProHisPheThrPheSerIleuGlySerIleGlnAsnAspTrpGluValSer 719
      1888 CCA---TTTAATTTGAACTGAGTAAAGCAATCTGTCACGAAAGTATGGAATCAAC 1944
      720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluIleValGly 739
      1945 AAGCTTAAACATCTATGATCCGAGTGTCTGCTTCAAA---AAGCTGAAAAGGCGCAAT 2001
      740 TyrValValIleuIleArgIleAsnAspGlyValArgProLeuGluGlyIleValSer 759
      2002 TACAAATCCCAATTCAGAGACAGATTTCTGGAAACACCTGTGACTAACACACAGAA 2061
      760 LeuProValThrPheCysSerCys 767
      2062 CTGAATTTACAGTGTGTTCTGC 2085

RESULT 13
US-08-474-067-1
; Sequence 1, Application US/08474067
; Patent No. 581518
; GENERAL INFORMATION:
; APPLICANT: Ranscht, Barbara
; TITLE OF INVENTION: T-Cadherin Adhesion Molecule
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,067
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/213,361
; FILING DATE: 14-MAY-1994

```

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/607,293
 FILING DATE: 30-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1682
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3959 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 45..2181
 US-08-474-067-1

Alignment Scores:
 Pred. No.: 6,22e-59 Length: 3959
 Score: 630.50 Matches: 193
 Percent Similarity: 44.77% Conservative: 115
 Best Local Similarity: 28.05% Mismatches: 283
 Query Match: 14.59% Indels: 97
 DB: 1 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-474-067-1 (1-3959)

QY 150 AlathrAspLeuAspAProAlaThrProAsnGlyGln-----LeuTYrTYrGlnIle 167
 DB 108 GCAGAAAGATTAGAA-----TGCACCCCTGGATTCCAGCAAAAGGTTTATT----- 155
 QY 168 ValIleGlnLeuProMetIleAsnAsnValMetTYrPheGlnIleAsnAsnIleTYrGly 187
 DB 156 ---ATTGAACAGCCA-----TTTGATTCCACAGAGACACCA 191
 QY 188 AlaIleSerLeuThrArgGlu-----GlySerGlnGluLeuAsn---ProAlaLys 203
 DB 192 ATTCGACCACTGGTGTGGTGGATGCTGCAGAGGGGAATTAACAATTGAACCTTGAAGTTTCT 251
 QY 204 AspProSerTYrAsnLeu-----ValIleSerValLysAspMetGly--- 217
 DB 252 AACCCAGACTTAAAGTGGAAACAGATGATCTTAACTGATGCAAGAAAGATATCAAA 311
 QY 218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
 DB 312 GCTGGCAGAGCTTTGTTGTCATGCACCGCTGAGCATGCT---GAGGATATGCGAGA 368
 QY 230 ValAspIleIleValIleThrGluAsnIleTYrLysAlaProLys----- 243
 DB 369 ATTTGATTGTTGGAGCTGATGAGAGACGATGCAATTAAGAAATCTTTAAGATGAA 428
 QY 244 -----ProValGluMetVal 248
 DB 429 GGCACCTTGGAAATTCAGACAAAAAGGCTATTCTGGCGACTCCAAATTTATATCCA 488
 QY 249 GluAsnSerThrAspProIleAspIleLysIleThrGlnValArgTYrAspAsp----- 266
 DB 489 GAAATATCAAGACACACATTTCCAGATCACTTGGCAAGTCATCAGAGAGGAAGGACA 548
 QY 267 ProGlyAlaGlnTYrSerLeuValAspLys-----GluLysLeuProAspPheProPhe 284
 DB 549 GAGGAGACCAAGTTCCGACTCTCTGTTAGGAGAGATCAAGACCCGAAAGGAATTTT 608
 QY 285 SerIleAspGln---GluGlyAspIleTYrValIleThrGlnProLeuAspArgGluGlnLys 303
 DB 609 AGAATTCATGATGATCACTGGGAGATGCTCTGTGACCCGACCCCTGGAGATGAGAAACATA 668
 QY 304 AspAlaTYrValIlePheTYrAlaValAlaLysAspGlnTYrGlyLysProLeuSerTYrPro 323

DB 669 GCCAATTATAGCTGGAGTTGAAGTAAACGGAATTTAAGTGGAAAAATCATGATGAGCCCA 728
 QY 324 LeuGluIleIleValLysValLysAspIleAsnAspAspProProThrCysProSerPro 343
 DB 729 GTCCGCTTGAATTTCTGTATTGATCAAAATGATTAACAGCCGATTTCAAAAGAGA 788
 QY 344 ValIleThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThr 363
 DB 789 CCTATGTTGGTCACTGATGAGAGATCCCTCAAGAAACAATGTATGCGGATGACA 848
 QY 364 AlaIleAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTYrArgIleValGlu 383
 DB 849 GCATTTGATGCTGATATATCTTACACAGCAACGCTCTTCTGGGTAAATCACTCAAG 908
 QY 384 GlnThrProLysLeuProMetAspGlyLeuPheLeuIleGln----- 397
 DB 909 CAGACACCTTACCAAACTTCCCAATATGTTCTTACATTGACCCAGAAAAAGGAGATAT 968
 QY 398 ---ThrTYrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLysGlnAspThrPro 416
 DB 969 GTCAAGTGGTGTCACTGTATCTGATCTGATCTGATGACATG-----GAAACGCCG 1019
 QY 417 GlnTYrAsnLeuThrIleGluValSerAspLys-----AspPheLysThr 431
 DB 1020 AAGTACGAGCTGCTTATTTGAAGCCAAAGATATGCGCCCATGATGAGACTTACTGCA 1079
 QY 432 LeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
 DB 1080 ACTGCAACTGCCACTATCTTATTTGATGACAAAAGACCAACCCACAGAAATTTACCAAG 1139
 QY 452 SerAspTYrGlyAsnLeuThrLeuValAspThrAsnIleGlySerThrIleLeuThr 471
 DB 1140 AAGAGATTT---CAGGCCACAGTAAGAGAA-----GAGTCAACAGAGTAATAGTAAC 1190
 QY 472 IleGlnAlaThrAspAlaAspGluProPheThrGlySerSerIleLysIleLeuTYrHisIle 491
 DB 1191 TTAACTGTTGTGACCGAGATGACCCAGCAACTGCGAGCATGAGACTGTCTACACTATT 1250
 QY 492 IleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProIleThrAsnThrGly 511
 DB 1251 ATTAACGAAATTCACAGGAGAGCTTTGAATTCATACATCCAGACTATATGAGGA 1310
 QY 512 TYrValIleIleLysIleAspProLeuAspPheGluThrAlaAlaValSerAsnIleValPhe 531
 DB 1311 ATGCTCTGTGTTCTCAACCTTATGACTATGAGATTTCAGCATTCACACATTCGTGATA 1370
 QY 532 LysAlaGluAsnProGluProLeuValPheGlyValLysTYrAsnAlaSerSerPheAla 551
 DB 1371 AAAGTAAAGAAATGAAACCCGTTGATTCCAGACATGACCTGACGCTCCAGTTCCACAGCA 1430
 QY 552 LysPheThrLeuIleValIleThrAspValAsnGluAlaProGlnPheSerGlnHisValPhe 571
 DB 1431 ACAGTTCAAGTCAACCGTTGAGATGTAAGAGCCCTGTTTCCACCCAAACCAATG 1490
 QY 572 GlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLys 591
 DB 1491 ACAGTGAACAAAACAAAGAAACATCCATTTGGAGCATTTGTGTAACAGTAATGCACT 1550
 QY 592 AspProGluGlyLeuAsp-----IleSerTYrSerLeuArgGlyAspThrArgGly 608
 DB 1551 GATCCAGACTCTTGGCAACATCAGACTATCAGGTATTCAGTTTCAAGGATCCAGCAAGC 1610
 QY 609 ThrLeuLysIleLeuAspIleValIleThrGlyGluIlePheSerValAlaProLeuAspArgGlu 628
 DB 1611 TGGCTAGAGTAATATCCACCAATGATGACCTGTCACCACTGCTCTGATCGGGAA 1670
 QY 629 Ala-----GlySerProTYrArgValGlnValAlaIleThrGluValGlyGly 644
 DB 1671 TCTCTCATGTTCAAGATTAACAATACACTGCTCTTCTCTGCAATAGCAAGTGTAAAC 1730
 QY 645 SerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAspPro 664
 DB 1731 CTTCTGTCTACAGTACAGAACTTAAACATCATCACCCTTGAGAGACGTCAATGACAAATGTC 1790

QY 665 -----ProAlaGluAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 1791 CCCTCCCTTACCAACACTGGCAAAA-----GCTCGATGATGATGCT 1832
 QY 681 SerIaProGlySerLeuLeuIlePheGluAlaThrAspAspArgHisLeuPheArgGly 700
 Db 1833 AAGATCTCAGAGTGGTACTAGAGCATCAGACAAAGCCCTCCATCCCAACACAGAT 1892
 QY 701 ProIlePheThrPheSerLeuGlySerGlySer---LeuGlnAsnAspTyrGluValSer 719
 Db 1893 CCA---TTTAATTGTAAGTGAAGTGAACATCTGCTCCAGAAAGTTATGAGATGATC 1949
 QY 720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAla 739
 Db 1950 AAGCTTAACATACTCATGCGCCAGGTGTCTGCTTCA---AACCTGAAAAAGCCCAT 2006
 QY 740 TyrValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSer 759
 Db 2007 TACAAACATCCCAATCTCAGTGAACAGATCTTGAAACACCACTTGACTTAACAACAGAA 2066
 QY 760 LeuProValThrPheCysSerCys 767
 Db 2067 CTGAATTACAGATGCTGTCTCTGC 2090

RESULT 14

US-08-474-068A-1
 ; Sequence 1, Application US/08474068A
 ; Patent No. 5837525

GENERAL INFORMATION:

APPLICANT: Ranscht, Barbara
 TITLE OF INVENTION: T-Cadherin Adhesion Molecule
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: United States
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/474,068A
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/213,361
 FILING DATE: 14-MAY-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/607,293
 FILING DATE: 30-OCT-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Campbell, Cathryn A.
 REGISTRATION NUMBER: 31,815
 REFERENCE/DOCKET NUMBER: P-LJ 1683
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3959 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 45..2181
 US-08-474-068A-1

Alignment Scores:
 Pred. No.: 6,22e-59
 Score: 630.50
 Percent Similarity: 44.77%
 Best Local Similarity: 28.05%
 Query Match: 14.59%
 DB: 2
 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-474-068A-1 (1-3959)

QY 150 AlAlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTyrTyrGlnIle 167
 Db 108 GCAGAAAGATTAAAG-----TGACACCCCTGAGATTCCAGAAAGGTTTTTAT----- 155
 QY 168 ValIleGluLeuProMetIleAsnValMetTyrPheGlnIleAsnValSerThrGly 187
 Db 156 ---ATTGAACAGCCA-----TTTGAATTCACAGAGACCAAGCCA 191
 QY 188 AlaIleSerLeuThrArgIle-----GlySerGlnIleAsn---ProAlaLys 203
 Db 192 ATTCTGAACCTGGTGTGTTGATGACTGCACAGGGAATTAACAAATGAACTTCGAACTTC 251
 QY 204 AsnProSerTyrAsnLeu-----ValIleSerValLysAspMetGly--- 217
 Db 252 AACCCAGACTTTAAGGTGAACACAGATGATCTTTAGTTGCACTGAAGAAATGATACGAA 311
 QY 218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
 Db 312 GCTGGCAGACGCTTTGTTTGTCCATGCACGCTGTGACAGATCT---GAGATATGACAGAA 368
 QY 230 ValAspIleIleValThrGluAsnIleTyrIleAlaProLys----- 243
 Db 369 ATTTGATTGTGAGCTGATGAGACAGACAGATGATTAAGAAATCTTTAAGATAGAA 428
 QY 244 -----ProValGluMetVal 248
 Db 429 GGCACCTTGAATTCACAGACAAAAAGGCTATTCTGCGCATCTCCAAATTAATTCAC 488
 QY 249 GluAsnSerThrAspProHisProIleLysIleThrGlnValArgTyrAsnAsp----- 266
 Db 489 GAAATATCAAAAGACACCACTATTTCCAGATCAGATGGCAGAGTCAATCAGAGTGAAGGACA 548
 QY 267 ProGlyAlaGlnTyrSerLeuValAspLys-----GlnLysLeuProArgPheProPhe 284
 Db 549 GAGGAGCAAAAGTCCGACTCTCTGTAAGGAGTATGACAAAGCCCAAGAAATTTT 608
 QY 285 SerIleAspGln---GluGlyAspIleTyrValThrGlnProLeuAspArgGluGluLys 303
 Db 609 AGATCAATGATGATCAGTGGGAGTGTCTGTGACCCGACCCCTGGATAGAGAACAAAT 668
 QY 304 AspAlaTyrValPheTyrValValAlaValAspArgIleTyrGlyLeuProLeuSerTyrPro 323
 Db 669 GCCAATTATGAGCTGGAAGTTGAAGTACGGAATTAAGTGGAAATATGATGAGCCCA 728
 QY 324 LeuGluIleHisValLysValLysAspIleAsnAspAspProProThrCysProSerPro 343
 Db 729 GTCCGCTTGATATTTCTGTTATGATCAAAATGATTAAGGCCCATGTTCAAGAGAGA 788
 QY 344 ValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrLeuThr 363
 Db 789 CCTATGTTGTGACCTCATGAGAGATCCCTCAAGAAACATGTATGAGGATGACA 848
 QY 364 AlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrTyrGlyIleValGlu 383
 Db 849 GCATTTGATGCTGATGATCTTAGCACAGCAACGCTCTTTCGGGTATTAATCCTCAAG 908
 QY 384 GlnThrProLysLeuProMetAspGlyLeuPheLeuIleGln----- 397
 Db 909 CAGACACCTACCAACCTTCCCAAAATATGTTCTACATTTGACCCCAAGAAAGGAGATATT 968
 QY 398 ---ThrTyrAlaGlyMetLeuGlnLeuValLysGlnSerLeuLysGlnAspThrPro 416
 Db 969 GTCACAGTGTGTCACTGTACTGCTGATCGTGAAGACATG-----GAAAGCGCG 1019

```

Qy 417 GlnTyrAsnLeuThrIleGluValSerAspLys-----AspPheLeuThr 431
Db 1020 AAGTACAGAGCTGTTATTAAGCCAGAGATATGCGCGCATATGTCGACCTTACTGGA 1079
Qy 432 LeuCyAspPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLys 451
Db 1080 ACTGCACTCCGACCTATTTCTTATTGATGACAAAACGACACCCACGACGATTTACCAAG 1139
Qy 452 SerAspTyrGlyAsnLeuThrLeuValGluAspThrAsnIleGlySerThrIleLeuThr 471
Db 1140 AAGGAGTTT---CAGGCCACAGATAAAGAA-----GAGGTCACAGAGATTAAGTAAAC 1190
Qy 472 IleGlnIleThrAspValAspGluProPheThrGlySerSerLysIleLeuThrIle 491
Db 1191 TTAACTGTGTGACGAGATGACCCGACCACTGAGACATGAGAGCTGTCTACACTATT 1250
Qy 492 IleGlyGlyAspSerGluGlyArgLeuGlyValAspThrAspProIleThrAsnThrGly 511
Db 1251 ATTAACGGAATCCAGGGCAGAGTTTGAATCATCATCAATCCCGACGATTAATGAGGA 1310
Qy 512 TyrValIleIleLysIleProLeuAspPheGluThrAlaValSerAsnIleValPhe 531
Db 1311 ATGCTCTGTGTGCAAACTTTAGACTATGATGATTCAGCATTTCCACATGCTGATATA 1370
Qy 532 LysAlaGluAsnProGluProLeuValPheGlyValLysTyrAsnAlaSerSerPheAla 551
Db 1371 AAAGTAAATAATGAAAGCCGTTATATTCAGACATACAGCTTACGCTCCAGTTCACAGCA 1430
Qy 552 LysPheThrLeuIleValThrAspValAsnGluValProGlnPheSerGlnIleValPhe 571
Db 1431 ACAGTTGATCATCCGTTGAGAGTGAATGAAGCCCTGTTTCCACCAAAACCAATG 1490
Qy 572 GlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaLys 591
Db 1491 ACAAGTCAAAACAGAGAAACATCCATTGCGCATTTGCGTAAACAGTAATATGCCACT 1550
Qy 592 AspProGluGlyLeuAsp-----IleSerTyrSerLeuArgGlyAspThrArgGly 608
Db 1551 GATCAGATACCTTGTGCAACATCAGACTATTCAGGATTCAGTTTCAAGATCCAGCAAC 1610
Qy 609 TrpLeuLysIleAspHisValThrGlyLysIlePheSerValAlaProLeuAspArgGlu 628
Db 1611 TGGCTAAGATTAATCCACCAATGATACGTCGACCACTGTCCTGTCGATCGGGA 1670
Qy 629 Ala-----GlySerProTyrArgValGlnValValAlaThrGluValGlyGly 644
Db 1671 TCTCTCATGTTTCAGATTAACAATACACTGCTCTCTCCGCGCAATAGACAGTGTATAC 1730
Qy 645 SerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAspPro 664
Db 1731 CTTCTCTCATGTTTCAGATTAACAATACACTGCTCTCTCCGCGCAATAGACAGTGTATAC 1790
Qy 665 -----ProArgLeuValAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
Db 1791 CCTCCTTATACCAACACTGCGAAA-----GCTGTGATGATGATGCT 1832
Qy 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspArgGlnIlePheAspArgGly 700
Db 1833 AAAGATCTCAGAGATGCTGCTACTAGAGCATCAGACAAAGCTCCATCCACACACAGAT 1892
Qy 701 ProHisPheThrPheSerLeuGlySerGlySer---LeuGlnAsnAspTyrGluValSer 719
Db 1893 CCA---TTTAATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1949
Qy 720 LysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAla 739
Db 1950 AAGCTTAACATATCTATGCTCCAGGTTCTCTGCTCA---AAGCTGAAAAGGCCAAT 2006
Qy 740 TyrValValLeuIleArgIleAsnAspGlyValArgProProLeuGluGlyIleValSer 759
Db 2007 TACAACATCCCAATCTCAGTACAGATTTCTGAAAACCACTCTGACTAACAACACAGAA 2066

```

```

Qy 760 LeuProValThrPheCysSerCys 767
Db 2067 CTGAATTAACAGTGTGTTCTCTGC 2090

RESULT 15
US-08-472-481-1
/ Sequence 1, Application US/08472481
/ Patent No. 5863804
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSER: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,481
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1686
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-8001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3959 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 45..2181
US-08-472-481-1

Alignment Scores:
Pred. No.: 6,22e-59 Length: 3959
Score: 630.50 Matches: 193
Percent Similarity: 44.77% Conservative: 115
Best Local Similarity: 28.05% Mismatches: 283
Query Match: 14.59% Indels: 97
DB: 2 Gaps: 25

US-10-025-380-1081 (1-832) x US-08-472-481-1 (1-3959)
Qy 150 AlaThrAspLeuAspAspProAlaThrProAsnGlyGln-----LeuTyrTyrGlnIle 167
Db 108 GCAGAGATTTAATA-----TGCACCCCTGCGATTCACAGCAAAAGTTTAT----- 155
Qy 168 ValIleGlnLeuProMetIleAsnAsnValMetTyrPheGlnIleAsnAsnValThrGly 197
Db 156 ---ATTGAACAGCA-----TTGAATTCACAGAGACAGCA 191
Qy 188 AlaIleSerLeuThrArgGlu-----GlySerGlnGluLeuAsn---ProAlaLys 203
Db 192 ATTCTGAACCTGCTGTTGATGATGACATGCAAGGGGAATTAACAAATTGAACCTTGAAGTTCT 251

```

QY 204 AsnProSerTyraAsnLeu-----ValIleSerValIysAspMetGly--- 217
 Db 252 AACCCAGACTTAAAGTGGAAACAGATGATCTTATTGATGACATGAAATGATACGAA 311
 QY 218 -----GlyGlnSerGluAsnSerPheSerAspThrThrSer 229
 Db 312 GCTGGCAGAGCTTGTGTTCATGACGCGTCTGACATGCT--GAGGATATGGCAGAA 368
 QY 230 ValAspIleIleValThrGluAsnIleTyrIysAlaProIys----- 243
 Db 369 ATTATGATTGTGAGCTGATGAGACAGCATCATTAAGAAATCTTAAGATACAA 428
 QY 244 -----ProValGluMetVal 248
 Db 429 GGCACCTGGAAATTCAGACAAAAAGGGCTATCTGGGCACTCCAAATATTAATTC 488
 QY 249 GluAsnSerThrAspProIleAspIleThrGlnValArgTyrAsnAsp----- 266
 Db 489 GAAATTCAGAACACCAATTCATCCAGATCAGTTGGCAGATCATCAGAGTGAAGGACA 548
 QY 267 ProGlyAlaGlnTyrSerLeuValAspIys-----GluIysLeuProArgPheProPhe 284
 Db 549 GAGGAGACAAAGTCCGACTCTCTGTAAAGGAGTACATCAGACCCGAAAGCAATTTT 608
 QY 285 SerIleAspGln--GluGlyAspIleTyrValThrGlnProLeuAspArgGluGluIys 303
 Db 609 AGATCAATAGATGATCAGTGGGATGTCTGTGACCCGACCCCTGGATAGAGAACATA 668
 QY 304 AspAlaTyrValPheTyrAlaValAlaIysAspGluTyrGlyIysProLeuSerTyrPro 323
 Db 669 GCCAATTATGAGCTGAGTGAAGTGAATTAAGTGGGAAATCAATGATGAGCCCA 728
 QY 324 LeuGluIleHisValIysValIysAspIleAsnAspAsnProProThrCysProSerPro 343
 Db 729 GTCCGCTTAAGATTTCTGTATTGATCAAAATGATTAACAGCCGATGTTCAAGAAAGA 788
 QY 344 ValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGlyThrThr 363
 Db 789 CCTATGTTGTGTCACGTCATGAGAAAGATCCCTCAACAGAACACTGTGATGCCGATACA 848
 QY 364 AlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArgIleValGlu 383
 Db 849 GCAATTGATCTGATGATCTGACACAGACAAACGCTCTGCGGTATTAACATCTCAAG 908
 QY 384 GlnThrProIysLeuProMetAspGlyLeuPheLeuIleGln----- 397
 Db 909 CAGACACTTACCAAACTTCCCAATATATGTTCACTTGAACCCAGAAAGGAGATATT 968
 QY 398 ---ThrTyrAlaGlyMetLeuGlnLeuAlaIysGlnSerLeuIysGlnAspThrPro 416
 Db 969 GTACACAGTGTGCACCTGCTACGCTGATGATCGTACAGCAATG-----GAAACGCC 1019
 QY 417 GlnTyrAsnLeuThrIleGluValSerAspIys-----AspPheIysThr 431
 Db 1020 AAGTACAGCTGCTGATTATTAAGCAAGATATGAGCGGCATGATGGGACTTACTGCA 1079
 QY 432 LeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluIys 451
 Db 1080 ACTGCAACTGCCACTATCTTATTGATGACAAAACGACCAACCAAGAAATTAACAG 1139
 QY 452 SerAspTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThr 471
 Db 1140 AAGGAGTT---CAGGCCACAGTAAAGAA-----GAGTCAACAGAGTAAATGATTAAC 1190
 QY 472 IleGlnAlaThrAspAlaAspGluProPheThrGlySerSerIysIleLeuTyrHisIle 491
 Db 1191 TTAACGTGTGATGACCAAGATGACCCAGCAACTGAGCATGAGAGCTGTCTACATATT 1250
 QY 492 IleIysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThrGly 511
 Db 1251 ATTAACGGAATTCAGGGCAGAGTTTGAATCATACCAATCCCGAGACTAATGAGGA 1310

QY 512 TyrValIleIleIysIysProLeuAspPheGluThrAlaIleValSerAsnIleValPhe 531
 Db 1311 ATGCTCTCTGTGCAAACTTTAGACTATGAGATTTACGATTTCCACATCTGCGATA 1370
 QY 532 LysAlaGluAsnProGluProLeuValPheGlyValIysTyrAsnAlaSerPheAla 551
 Db 1371 AAGTAGAAAATAGAACCCGTTGATTCAGACATGACCTACCGTCCCACTTCCACAGCA 1430
 QY 552 LysPheThrLeuIleValThrAspValAsnGluAlaProGlnPheSerGlnHisValPhe 571
 Db 1431 ACAGTTCAGATCCCGTTGAGATGTGATGAAGAGCCCTGTTTCCACCCAAACCAATG 1490
 QY 572 GlnAlaIysValSerGluAspValAlaIleGlyThrLysValGlyAsnValThrAlaIys 591
 Db 1491 ACAGTGAACAAACAGAGAAACATCCCTATTTGGCAGCATGTGTTAACATGAATGCACT 1550
 QY 592 AspProGluGlyLeuAsp-----IleSerTyrSerLeuArgIysThrArgGly 608
 Db 1551 GATCCAGATACCTTTGACATCAGACTATCAGGTATTCAGTTTACAGATCCAGCAGAC 1610
 QY 609 TyrLeuIysIleAspHisValThrGlyGluIlePheSerValAlaProLeuAspArgGlu 628
 Db 1611 TGCTTAGATTAATCCACCAATGATGATCCCTGTCACCACTGCTGTGATCGGGA 1670
 QY 629 Ala-----GlySerProTyrArgValGlnValAlaIleThrGluValGly 644
 Db 1671 TCCCTCATGTCAGATTAACAAATACATGCTCTCTCTGCAATGACATGATGAAC 1730
 QY 645 SerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspValAsnAspAsnPro 664
 Db 1731 CCTCTGCTACAGTACAGAACTTACATCATCCTTGGAGGACGTCATGACAAATGTC 1790
 QY 665 -----ProArgLeuAlaIysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db 1791 CCTCCCTTACCAACACTGGSCAAA-----GTCGTGATGATGCT 1832
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db 1833 AAGATCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1892
 QY 701 ProHisPheThrPheSerLeuGlySerGlySer---LeuGlnAsnAspTyrGluValSer 719
 Db 1893 CCA---TTTAATTTGAACTGAGTAAAGCAATCTGTCCAGAAAGTTATGAGAAATCAAC 1949
 QY 720 LysIleAspGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluAlaGly 739
 Db 1950 AAGCTTAACATATCTATGCGCAGGTGTCTGCTTCA--AACCTGAAAAGGCGCAAT 2006
 QY 740 TyrValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSer 759
 Db 2007 TACAAATCCCAATCTCAGTACAGATTTCTGGAACCAACCTTGTACTAACAACACAGAA 2066
 QY 760 LeuProValThrPheCysSerCys 767
 Db 2067 CTGAATTAACAAGTGTGCTCTGC 2090

Search completed: June 22, 2003, 03:24:10
 Job time : 217 secs

THIS PAGE BLANK (USPTO)

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2003, 21:24:29 ; Search time 62 Seconds
(without alignments)
1788.137 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MILDGHLHSLCLMLYLATG.....DKGDNVESAGASVYPLRS 8322

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```

Searched:      908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters:  908470

```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
_A_GeneSeq_101002:*
1: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/genseq/genseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/genseq/genseqp-emb1/AA2002.DAT:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4321	100.0	832	20	AAW74089	Human HPI-1 protein
2	4321	100.0	832	22	AAW24522	C880P similar amilin
3	908.5	21.0	807	21	AAV95405	Human PRO1340 (UNP
4	908.5	21.0	807	22	AAW87574	Human PRO1340. Hs
5	908.5	21.0	807	22	AAW66154	Protein of the inv
6	908.5	21.0	807	23	AAU81685	Human PRO protein.
7	877.5	20.3	830	22	AAU09959	Human cadherin-2 (
8	699.5	16.2	906	21	AAV70741	Human N-cadherin.
9	699.5	16.2	906	22	ABG1416	Novel human diagm
10	699.5	16.2	906	23	ABW8174	Human N-cadherin f

15	569.5	13.2	840	22	AAW23993	Human EST encoded
44	570.5	13.2	840	22	ABG01693	Novel human diaphn
43	571.5	13.2	796	22	AAH57254	Mouse ischaemic c
42	571.5	13.2	788	22	AAH78375	Human protein (SQ
41	573.5	13.3	796	15	AAH49730	Sequence encoded)
40	574.5	13.3	796	22	AAH60418	Mouse PTP-binding
39	576.5	13.3	3028	22	AAH08566	Human NOY7 protei
38	576	13.3	3014	22	AAU02166	Seven-pass transm
37	576	13.3	3014	22	AAU68533	Human novel cytol
36	591	13.7	882	22	AAU78051	Human E-cadherin
35	591	13.7	882	22	AAH73450	Human E-cadherin
34	591	13.7	882	21	AAH35730	Human E-cadherin
33	591	13.7	882	20	AAH09375	Wild-type E-cadhe
32	591.5	13.7	4643	22	ABH71609	Drosophila melano
31	592.5	13.7	117	21	AAH45597	Nonclassical cadh
30	592.5	13.7	117	21	AAH45597	Nonclassical cadh
29	593	13.7	878	16	AAH85487	Human E-cadherin
28	593	13.7	878	15	AAH55060	Sequence of human
27	601.5	13.9	3606	22	ABH52595	Drosophila melano
26	604	14.0	899	23	ABH81472	Mouse E-cadherin
25	605	14.0	899	23	ABH81472	Mouse E-cadherin
24	607	14.0	4591	22	ABG22977	Novel human diaphn
23	611	14.1	4555	22	AAH52106	Rat fat 3 protein
22	630.5	14.6	717	13	AAH27824	Sequence encoded)
21	630.5	14.6	712	13	AAH3136	Full length human
20	657.5	15.2	713	18	AAH25638	Human cadherin-13
19	657.5	15.2	862	22	AAH1021	Human polypeptide
18	665.5	15.4	862	22	AAH59235	Human desmocollin
17	667.5	15.5	847	22	AAH39255	Human polypeptide
16	671.5	15.5	896	16	AAH3533	Human HT-1376 cel
15	671.5	15.5	896	16	AAH3533	Human HT-1376 cel
14	684.5	15.8	916	18	AAH3129	Full length human
13	684.5	15.8	916	18	AAH35658	Human cadherin-4
12	694.5	16.1	916	23	ABH52233	Chicken N-cadheri
11	696	16.1	912	23	ABH84471	Chicken N-cadheri

ALIGNMENTS

RESULT	1
AAW74089	
ID	AAW74089 standard; Protein; 832 AA.
AC	
XX	AAW74089;
DT	04-MAY-1999 (first entry)
XX	
DE	Human HPT-1 protein sequence.
XX	
KM	Gastro-intestinal transport receptor; binding protein; hSI; HPT1;
KM	DH; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KM	intestinal peptide-associated transporter; hypertension; diabetes;
KM	osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KM	therapeutic agent delivery; therapy; probe.
OS	Homo sapiens.
XX	
PN	WO9851325-A2.
XX	
PD	19-NOV-1998.
XX	
PF	15-MAY-1998; 98WO-US10088.
XX	
PR	15-MAY-1997; 97US-0046595.
XX	
PA	(CYTO-) CYTOGEN CORP.
PA	(ELAN-) ELAN CORP PLC.
PI	Alvarez VL, Belinka BA, Cagney GW, Carter JM, Lambkin LJ;
PI	Onahony DJ, Patterson CA, Singleton J;
RR	WPI; 1999-009568/01.

DR N-PSDB; AAX18166.

XX New proteins that bind specifically to receptors in the

PT gastro-intestinal tract and related nucleic acid - chimeras and

PT antibodies, used to deliver therapeutic or diagnostic agents to, or

PT through, the gastrointestinal tract, e.g. insulin or leuprolide

XX

PS Disclosure; Fig 2; 294pp; English.

XX This sequence is the human HPT-1 protein. The invention relates to

CC purified proteins (I) that bind specifically to at least one of the

CC gastro-intestinal (GI) tract receptors human intestinal

CC peptide-associated transporter (HPT1), hPPT1, D2H and human

CC nucleoside-isomerase complex (hSI). (I) provide active transport of

CC therapeutic agents through human and animal GI tissue (into the blood)

CC for in vivo delivery, particularly for treatment or prevention

CC of hypertension, diabetes, osteoporosis, haemophilia, anaemia, cancer,

CC migraine, or angina pectoris. Specifically they are used to deliver

CC insulin or leuprolide, but many other suitable therapeutic agents are

CC disclosed, including genes or inhibitory nucleic acid, imaging agents and

CC antigens. (I) may also provide targeting to the GI tract. Other uses of

CC (I) are: (i) to determine the level of specified receptors in a sample

CC (in a binding assay); and (ii) to screen for molecules that bind (I).

CC Immunogenic analogues or derivatives of (I) are used to raise antibodies

CC and in immunoassays. The antibodies are used to locate, detect and

CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis

CC etc., also for peptide purification and immobilisation.

XX

XX Sequence 832 AA;

Query Match 100.0%; Score 4321; DB 20; Length 832;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MILQAHLSLCLMLYLATGYGQEGKFGSPKPMFTSYEGEGSPQIIFQKAPPAVTF 60

DB 1 MILQAHLSLCLMLYLATGYGQEGKFGSPKPMFTSYEGEGSPQIIFQKAPPAVTF 60

QY 61 ELTGETDNI FVIEREGGLLYNRALDRETRSTHNLQVALDANGIIIEGVPITTEVKDIN 120

DB 61 ELTGETDNI FVIEREGGLLYNRALDRETRSTHNLQVALDANGIIIEGVPITTEVKDIN 120

QY 121 DNRPTFLOSKEYGVRQNSRGKPELYNATDLDPAFNPQOLYYQIYIOLPMNNWYF 180

DB 121 DNRPTFLOSKEYGVRQNSRGKPELYNATDLDPAFNPQOLYYQIYIOLPMNNWYF 180

QY 181 QINNKTGAISLTRREGSCELNPAPKPSYNLVISVKMGQSGNSFSDTTSVDIIVTENIMK 240

DB 181 QINNKTGAISLTRREGSCELNPAPKPSYNLVISVKMGQSGNSFSDTTSVDIIVTENIMK 240

QY 241 APKPEVENENSTDPHPKITQVRANDPQAOYSLVDKELPPRPPSIDEGSIYYTQPLDR 300

DB 241 APKPEVENENSTDPHPKITQVRANDPQAOYSLVDKELPPRPPSIDEGSIYYTQPLDR 300

QY 301 EEKQAYVYAAKDEYGPISYPLEIHKVADINDNPPTCSPTVFEVQENERLGNISIG 360

DB 301 EEKQAYVYAAKDEYGPISYPLEIHKVADINDNPPTCSPTVFEVQENERLGNISIG 360

QY 361 TLTAHADRDEENTANSFLNRYIVEQTPKLPMDGLFLIQTYAGMLQIAKOSLKKDTPQYNL 420

DB 361 TLTAHADRDEENTANSFLNRYIVEQTPKLPMDGLFLIQTYAGMLQIAKOSLKKDTPQYNL 420

QY 421 TIEVSDQKFTLCPQVQVINDINOIPIFEKSDVGNLTLADTNIGSTILTIOATDDEP 480

DB 421 TIEVSDQKFTLCPQVQVINDINOIPIFEKSDVGNLTLADTNIGSTILTIOATDDEP 480

QY 481 FTGSSKILYHITIKGDSRGLGVDTPHNTNGVYIIKKPLDEETAANSIVKAEKNEPELV 540

DB 481 FTGSSKILYHITIKGDSRGLGVDTPHNTNGVYIIKKPLDEETAANSIVKAEKNEPELV 540

QY 541 FGKYNASSPAKFTLITVDVNEAPQFSGHVFQAKSEVVAIGTKKGNNTAODPGLDISY 600

DB 541 FGKYNASSPAKFTLITVDVNEAPQFSGHVFQAKSEVVAIGTKKGNNTAODPGLDISY 600

QY 601 SLRGDTRGMLKIDHVTGEIFSVAPLDREAGSPYRVQVAVATEVGGSSLSVSEPHLLMDV 660

DB 601 SLRGDTRGMLKIDHVTGEIFSVAPLDREAGSPYRVQVAVATEVGGSSLSVSEPHLLMDV 660

QY 661 NDNPPRLAKDYTGFLFCHPPLSAGSLIFEAITDDQHLFRGPHFTFSLGSSSLONDEWVSK 720

DB 661 NDNPPRLAKDYTGFLFCHPPLSAGSLIFEAITDDQHLFRGPHFTFSLGSSSLONDEWVSK 720

QY 721 INGTHARLSRHRDPEERAVVVLIRINDGRRPLEGIVSLPTFCSCVEGSCFRPAGHOT 780

DB 721 INGTHARLSRHRDPEERAVVVLIRINDGRRPLEGIVSLPTFCSCVEGSCFRPAGHOT 780

QY 781 GIPFVGAVGILLTTLVLIGIILAVVFIKIKDKGKNVESAQASEVKPIRS 832

DB 781 GIPFVGAVGILLTTLVLIGIILAVVFIKIKDKGKNVESAQASEVKPIRS 832

RESULT 2

AAM24522

ID AAM24522 standard; Protein; 832 AA.

XX

AC AAM24522;

DT 12-OCT-2001 (first entry)

XX

DE C880P similar amino acid sequence (GENESEQ W740898).

XX

KM Human; immunotherapy; diagnosis; colon cancer; colon tumour;

KW immunogenic; gene therapy; vaccine; colonic cancer.

XX

OS Homo sapiens.

XX

EN WO200149716-A2.

XX

PD 12-JUL-2001.

XX

PF 29-DEC-2000; 2000MO-US35596.

XX

PR 30-DEC-1999; 99US-0476296.

PR 10-JAN-2000; 2000US-0480321.

PR 15-FEB-2000; 2000US-0504629.

PR 06-MAR-2000; 2000US-0519444.

PR 19-MAY-2000; 2000US-0575251.

PR 29-JUN-2000; 2000US-0609448.

PR 28-AUG-2000; 2000US-0649811.

XX

PA (CORI-) CORIXA CORP.

XX

PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

XX

DR WPI; 2001-441847/47.

XX

PT Colon tumor associated proteins and nucleic acids useful for the

PT prevention, diagnosis and treatment of colonic cancer -

XX

PS Claim 2; Page 469-472; 472pp; English.

XX

CC The present invention describes colon tumour associated proteins (I) and

CC the polynucleotides (II) that encode them. (I) have cytostatic activity.

CC (I) and (II) can be used in gene therapy and vaccine production. (I) and

CC (II) may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate colon tumour associated protein (TCAP)

CC expression, such as colonic cancer. For example, (I) and (II) may be

CC used to treat disorders associated with decreased expression by

CC rectifying mutations or deletions in a patient's genome that affect the

CC activity of TCAPs by expressing inactive proteins or to supplement the

CC patients own production of them. Additionally, (II) may be used to

CC produce the TCAP proteins, by inserting the nucleic acids into a host

CC cell culturing the cell to express the protein. (II) and its

CC complementary sequences may also be used as DNA probes in diagnostic

CC polymerase chain reaction (PCR) and hybridisation assays to detect and

quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. (1) may also be used as antigens in the production of antibodies against TCAPs and in assays to identify modulators of TCAP expression and activity. Anti-(1) antibodies and antagonists may also be used to down regulate TCAP expression and activity. The anti-(1) antibodies may also be used as diagnostic agents for detecting the presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). AA128460 to AA129512 and AA24494 to AA24523 represent nucleotide and amino acid sequences given in the exemplification of the present invention.

XX Sequence 832 AA;

Query Match 100.0%; Score 4321; DB 22; Length 832;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MIIQAHHSICLMLYATGCGEGKSPGPKMTFSIYGOEPSQIIQPKANPAVTF 60
DB 1 MIIQAHHSICLMLYATGCGEGKSPGPKMTFSIYGOEPSQIIQPKANPAVTF 60
QY 61 ELTGETDNIVIREBGLLYNRALDRETRSTHNLQVALANGIIVGVPITTEVNDIN 120
DB 61 ELTGETDNIVIREBGLLYNRALDRETRSTHNLQVALANGIIVGVPITTEVNDIN 120
QY 121 DNRPTFLOSKEYSVRQNSRPGKFLVYNAITDLPATPGOLYYQIVILPMINMYF 180
DB 121 DNRPTFLOSKEYSVRQNSRPGKFLVYNAITDLPATPGOLYYQIVILPMINMYF 180
QY 181 QINNKKGALSLTBEGSGLNPAPKPSYNIYISVKMGSGSSESDTSDIIVTENTWK 240
DB 181 QINNKKGALSLTBEGSGLNPAPKPSYNIYISVKMGSGSSESDTSDIIVTENTWK 240
QY 241 APKPEVENSTPPIPKITQVRMNDGAPQSLVDEKELRFPSPSIDOEGDIIVTQPLDR 300
DB 241 APKPEVENSTPPIPKITQVRMNDGAPQSLVDEKELRFPSPSIDOEGDIIVTQPLDR 300
QY 301 EEKDAVYFVAADDEYKPLSYPLEIHVKYKINDNPPTCSPVTVFVQENRGLNSIG 360
DB 301 EEKDAVYFVAADDEYKPLSYPLEIHVKYKINDNPPTCSPVTVFVQENRGLNSIG 360
QY 361 TLTAHRDEENTANSFLNRYIVEQTPKLPMDGLFLIQTVAQMLQAKSLKKODTPQYL 420
DB 361 TLTAHRDEENTANSFLNRYIVEQTPKLPMDGLFLIQTVAQMLQAKSLKKODTPQYL 420
QY 421 TIEVSDKDEKTLCEVOINVINDIQPIPEKSDYGNLTAAEDNIGSTIITTIQATDAEP 480
DB 421 TIEVSDKDEKTLCEVOINVINDIQPIPEKSDYGNLTAAEDNIGSTIITTIQATDAEP 480
QY 481 FTGSSKILYHIIGDSEGRIGVDTDPHTNTGYVLIKPLDFTAAVNIYFKAENPELV 540
DB 481 FTGSSKILYHIIGDSEGRIGVDTDPHTNTGYVLIKPLDFTAAVNIYFKAENPELV 540
QY 541 FGKYNASSFAKFTLLIVTDVNEAPQPSQHVFOAKVSEDAVIGTKVGNVTAKDEGLDISY 600
DB 541 FGKYNASSFAKFTLLIVTDVNEAPQPSQHVFOAKVSEDAVIGTKVGNVTAKDEGLDISY 600
QY 601 SLRGDTRGMKIDHVTGEIYFSVAPLDREAGSPYRVQVVAITEVGGSSLSYSEPHLLIMDY 660
DB 601 SLRGDTRGMKIDHVTGEIYFSVAPLDREAGSPYRVQVVAITEVGGSSLSYSEPHLLIMDY 660
QY 661 NDNPPLAKDYTGALFCHPLSAPGSLIFEATDDOHLFRPHFTFSIGSGLONDNVSK 720
DB 661 NDNPPLAKDYTGALFCHPLSAPGSLIFEATDDOHLFRPHFTFSIGSGLONDNVSK 720
QY 721 INGTARLSTRATDFEBRAYVVLIRINDGSRPLEGIVSLVTFCSCEBSCRRPAHQOT 780
DB 721 INGTARLSTRATDFEBRAYVVLIRINDGSRPLEGIVSLVTFCSCEBSCRRPAHQOT 780
QY 781 GIPYGAAGIILTLTLLVGIILAVFIRIKKDKGDNVSAQSEKPKRS 832
DB 781 GIPYGAAGIILTLTLLVGIILAVFIRIKKDKGDNVSAQSEKPKRS 832

```

RESULT 3

AA199405 standard; Protein; 807 AA.

AA199405;

08-AUG-2000 (first entry)

Human PRO1340 (UNQ695) amino acid sequence SEQ ID NO:229.

Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;

transmembrane; secretion; immunoadhesion; pharmaceutical; screening.

Homo sapiens.

WO200012708-A2.

09-MAR-2000.

01-SEP-1999;

99MO-US20111.

01-SEP-1998;

98US-0098716.

01-SEP-1998;

98US-0098749.

01-SEP-1998;

98US-0098750.

02-SEP-1998;

98US-0098803.

02-SEP-1998;

98US-0098821.

02-SEP-1998;

98US-0098843.

02-SEP-1998;

98US-0098936.

02-SEP-1998;

98US-0098959.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

02-SEP-1998;

98US-0098962.

[illegible]

XX	Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
PI	WPI; 2000-237871/20.
XX	N-PSDB; AAA37087.
DR	New mammalian DNA sequences encoding transmembrane, receptor or
XX	secreted PRO polypeptides, useful for screening of potential peptide or
PT	small molecule inhibitors of the relevant receptor/Ligand interactions
XX	Claim 12; Fig 132; 773bp; English.
PS	AAA37022 to AAA37144 encode the new isolated human transmembrane,
CC	receptor or secreted PRO polypeptide given in AA99340 to AA99462. The
XX	transmembrane and receptor PRO proteins can be used for screening of
CC	potential peptide or small molecule inhibitors of the relevant
CC	receptor/Ligand interactions. The polypeptides and nucleotide sequences
CC	encoding them have various industrial applications, including uses as
CC	pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
CC	PCR primers and hybridization probes used in the isolation of the PRO
CC	polypeptides from the present invention.
XX	Sequence 807 AA;
SQ	
Query March	21.0%; Score 908.5; DB 21; Length 807;
Best Local Similarity	30.2%; Pred. No. 1,Je-64;
Matchess 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21	
62	L TGE---TDNI FVIERE-GLLYNNRRLDRETRSTHNLQAALDANGIIEGVPVITTEV 116
Db	LSGSCKKTEBPAMDPSGLVTBALDREGAEYQLQVLTLEMDDGHVLWGCPVLVHV 116
Oy	117 KDINDNRPFLQSKYESGVRONSRPKFPFLYNATDLDPATPNGOLYYOYLQPLMTNN 176
Db	117 KDENDQVHFQAIYRAILSRGTBPGIPFLFLEASDRDEPTANSDFHLISOAPAPS 176
Oy	177 VMYFOINKKTALISLRESGELNDPAKNPSYLVIVSYDMGGQSNSPDSPTSVDIIYTE 236
Db	177 PDMFLEPRICATALSPKGSTSLBALERTTQLQVONKMDQDA-SHQAIRTVESVITE 235
Oy	237 NIKAKPEVEVENSTDPHPKITOVNRNMDSGAQYSLVDKETLRPFPSIDOGEDIVYTQ 296
Db	236 STWSLEBIHLAENLKVLVPHHMQLVHMSGGDVHNHL--ESHPRGPFEVNAEGMLVYTR 292
Oy	297 PLDREKDAYVFVAADKEYGKPLSYPLEIHVKYKDINDNPRTCSRYTFVEVGQNEKLQ 356
Db	293 ELDRBAQAYLIQVAAQNSHGEDVAAAPDELHYLVMDENDNVNICPPRDPTVISIELSPBG 352
Oy	357 NSIGTLTHDDEENTANSPFNRYRVBDQTPKLPMDG-LFLIQTYAGMQLAKOSLKK-OD 414
Db	353 TEVTRLSEADADAPGPSNHVVYQLSPEPEDGEGRAFOVDPTSGSVTLGYLPLRAGON 412
Oy	415 TPQVYLTLEVSDKD--FKTLFCVOINVINDINDQPIEFKSPBYGNLTLLAEDTNIGSTILLTI 472
Db	413 ILLVLANDLAGABEGFSSTCEVEVAAYDINDHAPEFTSQIGPISLEDVEPGLVAML 472
Oy	473 QATDAD-EPTFGSSKIIL-YHIIKGBSEGRLGVDDPTHTNTGYVIK--KPLDEFTRAASN 528
Db	473 TALDADLER--AFRLMDFAIRBEDGTGEGFDWEP--DSGHVRLRLCKNLSYEAPASHHE 527
Oy	529 IVFKANEDEPLFYGVKYNAASSFAKTTLYTDVYNAERPOSCNYFOAKVSEDAVALGRKVGV 588
Db	528 VVVVVQSAVKLY-GGPGPGATATATYTLVERVMPPKLDQSEYSAVSPIADAGSFLLTI 586
Oy	589 TAKDEGDIISYLSMGDTRGMKLIKDHVGETIFSVAPLD-REAGSPYRQOVVATEVSGSSL 647
Db	587 OPSDISSTLRSLVNDSBGMCLIEKFSGEVHTAOSLGAGRGDTYYTLVLEQDTR----642
Oy	648 SSVSEFHILMDVNDNPRLAKDYTGLEFCHPLSPAQSILF-EATDDOHLFRGPHFTFS 706
Db	643 -----LTLAPVPSQ-----YLCSTRQOHGIIVSGPSKDPDLASGHGP-YSPFT 683
Oy	707 LGSGLONDNEVSKIINGTHARLASLRHDFFERRAVVULIRINDGGRPRLIEGIVSLPVYTC 765

```

Db      684 LGNPTVQQRDMRQLTGLNGSHAYLTTLALHWVEPREHIIIPVAVSHNAQ---MMQLLVRYIVC 740
Qy      766 SC-VEGSCFRPAGHOTGIPFVGNVAVGIIITLTVIGIILAVFIRIKDKGKNVESAOA 824
Db      741 RCNVEGQCMKRGKVMKMPITLSAVGILVGTVAIGIFILIFHTMTMSRKDDPDQPADS 800
Qy      825 SEVK 828
Db      801 VPLK 804

RESULT 4
AAB87574
ID      AAB87574 standard; Protein; 807 AA.
AC      AAB87574;
XX
DT      15-MAY-2001 (first entry)
DE      Human PRO1340.
XX
KM      Human; PRO protein; mapping.
XX      Homo sapiens.
XX      MO200116318-A2.
XX      08-MAR-2001.
XX      24-AUG-2000; 2000MO-US233328.
XX      01-SEP-1999; 99MO-US20111.
XX      15-SEP-1999; 99MO-US21090.
XX      07-DEC-1999; 99US-0169485.
XX      09-DEC-1999; 99US-0170282.
XX      11-JAN-2000; 2000US-0175481.
XX      18-FEB-2000; 2000MO-US04341.
XX      22-FEB-2000; 2000MO-US04342.
XX      01-MAR-2000; 2000MO-US05601.
XX      03-MAR-2000; 2000US-0187202.
XX      25-APR-2000; 2000US-0193937.
XX      22-MAY-2000; 2000MO-US14042.
XX      05-JUN-2000; 2000US-0209832.
XX
PA      (GETH ) GENENTECH INC.
XX      Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ,
PI      Grimaldi CJ, Gurney AL, Watanabe CK, Wood WI;
XX
DR      WPI; 2001-183260/18.
XX      N-PSDB; AAP92106.
XX
PT      Eighty four nucleic acids encoding PRO polypeptides, useful in
PT      molecular biology, including use as hybridization probes, and in
XX      chromosome and gene mapping.
XX
PS      Claim 12; Fig 98; 278pp; English.
XX
CC      The present sequence is a human PRO polypeptide (secreted and
CC      transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC      anti-PRO antibodies are useful for preparation of a medicament useful in
CC      the treatment of a condition which is responsive to the PRO protein,
CC      agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC      employed as molecular weight markers for protein electrophoresis. The PRO
CC      coding sequence has applications in molecular biology, including use as
CC      hybridisation probes, and in chromosome and gene mapping.
XX
SO      Sequence 807 AA;
XX
Query Match 21.0%; Score 908.5; DB 22; Length 807;
Best Local Similarity 30.2%; Pred. No. 1.le-64;

```

```

Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
Qy      62 LTGE-----FDNIIVIRE-GLLYNRALDRETSSTNLOVAALDANGIIVEGVPTIEV 116
Db      57 LSGDSGKATBGPAMPDPSGFLVYTRALDREDAEYQLOVLTLEMDSHVIMGPOVIVVH 116
Qy      117 KDINDNRPTFLQSGKYGSVRQNSRPGKPELYVNATLDDPATNGOLYYOIVLOPMINN 176
Db      117 KDENDQVPHFSQAIYRARLSRGTRPGIFLFLASRDEGTANSDLRPHILLSQAAPQPS 176
Qy      177 VMTFQINNKTAISLITREGSDELPAKNSYNLVISYKMGGSSENSFSDTSVDIIVTE 236
Db      177 PDWFOLEPRIGALALPKSTSLDHALERTYQLLVQVKMGDQA-SGHQATATVEVSIIE 235
Qy      237 NIKAKPPEVMEVNSTDPHPKIKITQVRMNDPGAQYSIVDEKSLPRPFSDIDOGSDIYVTO 236
Db      236 STVSLSEPIHLAENLKVLPHMAQVHWGSGDVVHLL---ESHPPGFVEVNAEGNLVYTR 232
Qy      297 PLDREKDAVYFAVAKDEYKPLSYPLEIHVKDKINDNPPTCPSPVTVFEVQENERLG 356
Db      293 ELDRQAQAEYLLQVRAQNSHGEDIYAAPLELHVLMENDNVPCPPRPDPVSIPELSPPG 352
Qy      357 NSIGITLTAHDRENTANSFLNRYIVEQTPKLEMDG-LFLIQTYAGMLQAKOSLKK-QD 414
Db      353 TEVTRLSAEDADAPGSPNSHVYQLSPSEDEVEGRACVDPSTSGSVTLGVPLRAGON 412
Qy      415 TPQYNLTIEVSDKD--FKTLCFQVINVIDINDQIPIFEKSDVGNLTAAEDTNGSTLTI 472
Db      413 ILLVLAMDLAGAGGFSSTCEVEVAVTDINDAPFITSQIPISLPEDVEGTIVAML 472
Qy      473 QATDAD-EPFTGSSKIL-YHIKDGSEGLGVDDPTHTVGVYIK--XPLDETAVASN 528
Db      473 TATDADLEP---AFRLMDFAIERGDTGFTGLDMEF--DSGHVRLKCNLSYEAPSHS 527
Qy      529 IYFKAENPEPLVGVGKYNASSFAKFTLIYTDVNEAQQFQOHVQOAVSEDAVIGTKVNV 588
Db      528 VVVVQSVAKLV-GPSPGCGATATVTLVERVWPPPKLQOESYEASVPI SAPAGSFLTI 586
Qy      589 TAKDPGGLDISYLRDTRGMLKIDHVTGEIPEVAFLD-REAGSPRYOVVATEVGGSSL 647
Db      587 QPSDPIKRTIRPSLVNDSGMLCTIEKFSGVHTAOSLQAQPEDITYTLVLEADDTA---- 642
Qy      648 SSVSEFHLILMDVNDNPPRLAKDYTLFFCHPLSAPGSLIF-EATDDQHLFRGPHFTFS 706
Db      643 -----LTLAPVPSQ-----YICTPQDHLGVSGSKPDLASHGCP-YSTFT 683
Qy      707 LG-SGSLQNDMEVSKINGTHARLSTRHTDPEERAVVYLIRINDGSRPLEGIYSLPVTF 765
Db      684 LGNPTVQQRDMRQLTGLNGSHAYLTTLALHWVEPREHIIIPVAVSHNAQ---MMQLLVRYIVC 740
Qy      766 SC-VEGSCFRPAGHOTGIPFVGNVAVGIIITLTVIGIILAVFIRIKDKGKNVESAOA 824
Db      741 RCNVEGQCMKRGKVMKMPITLSAVGILVGTVAIGIFILIFHTMTMSRKDDPDQPADS 800
Qy      825 SEVK 828
Db      801 VPLK 804

RESULT 5
AAB66154
ID      AAB66154 standard; protein; 807 AA.
XX
AC      AAB66154;
XX
DT      02-APR-2001 (first entry)
DE      Protein of the invention #66.
XX
KM      Secreted; transmembrane; gene therapy.
XX
OS      Unidentified.
XX

```

PN WO200078961-A1.
 XX 28-DEC-2000.
 PD 18-FEB-2000; 2000WO-US04342.
 XX 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 01-SEP-1999; 99WO-US20111.
 PR 29-OCT-1999; 99US-0162506.
 PR 30-NOV-1999; 99WO-US28313.
 PR 02-DEC-1999; 99WO-US28551.
 PR 16-DEC-1999; 99WO-US30095.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00376.
 PA (GETH) GENENTECH INC.
 XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong SJ,
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ,
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;
 PI Watanabe CK, Williams PM, Wood WI;
 XX WPI, 2001-071395/08.
 DR Secreted and transmembrane proteins and nucleic acids designated PRO,
 PT useful as hybridization probes, in chromosome and gene mapping and gene
 therapy -
 XX
 PS Claim 1; Fig 132; 787pp; English.
 CC The present invention relates to secreted and transmembrane proteins.
 CC These proteins and the DNA encoding them may be used as hybridization
 CC probes, in chromosome and gene mapping and in the generation of
 CC anti-sense RNA and DNA. They may also be used to generate either
 CC transgenic animals or knockout animals which are in turn useful for
 CC development and screening of therapeutically useful reagents.
 CC The nucleic acids may also be used in gene therapy.
 CC
 XX
 SQ Sequence 807 AA;
 Query Match 21.0%; Score 908.5; DB 22; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.1e-64;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
 QY 62 LTGE---TDNIFVIERE-GLLYNRALDRETRSTHNLQVAAIDANGIIVEGPPVITTEV 116
 DB 57 LSGDSKATGSGPFAMDPDSGFLVTRALDREEQAEYQVTLFMDQGHVLMGPPVIAVHV 116
 QY 117 KDINDNRPTPLQSKYEGSVQRNSRPGRFELYVNAITDDDPATPPGOLYQYQIVQLPMINN 176
 DB 117 KDENDQVPHPSQALYRARLSRGTRPGIFFLFLASDRDEPTANSDLRFHLLSQAPQPS 176
 QY 177 VMYFOINNKTGALSLTREGSOELNPAKNPSYNLVYSVMKMGQSENSFSDTTSVDIIVTE 236
 DB 177 PDNFQLEPRIGALSLSPKSGTSLDHALERTYQLVQVQKMGDQA-SGHQATATVEGSIIE 235
 QY 237 NIWKAPEVEMVENS TDHPRIKITQVRWMDPQAQYSLVDEKELRPFPFSDIDQESDIYVQ 296
 DB 236 STWVSLERIHAEHLKYLVRPHMAQVHMSGGDVYHLL--ESHPPGFEEVAAENLVYTR 292
 QY 297 PLDEEKDAYFYVAVADEYKPLSTYPLEIHVKKYKDINDNRPTCPSPYTVVEVQENRIG 356
 DB 293 ELDEEAQAEYLLQVRAQNSHGEDYAAPLEHLVLVWDENDVNPICPPRPPTVSIELSPPG 352
 QY 357 NSIGTLAHDDEENTANSFLNRYIVEQTPKLPMDG-LFLQTYAGMLQAKQSLKK-OD 414
 DB 353 TEVTRLSHEDADAGSPNSHVYQLSPBEDGVGBRAFYQVDPPTSGSVTLGVLELRAGQN 412
 QY 415 TPQYNLTIEVSKD--FKTLCFVQIVINDIQIPIFEKSDYGNMLTAEDTNIGSTILTI 472
 DB 413 ILLVLAMDLAGAGGFSSTCEVEVAVTIDINDHAPEFITSGIGISLPEDEVPEGTIVAML 472

QY 473 QATDAD-BPFTGSSKIL-VHIKGDSEGRLGVDTPDPTNTNGYIIR--KPLDETAVERN 528
 DB 473 TAIDADLEP--AFRLMDPAIERGDTGFTGLDMEP--DSGHVRLALCNLSYEAPRSH 527
 QY 529 IYFKANPEPLVFGVYKNASFPKFTLLIVTDVNEAPQSOHVFOAKYSDVAIGTRVGNV 588
 DB 528 VVVVYQSVAKLV-GPEPGQATFATVTLVERVWPPKLDQESYEASVPI SAPAGSFLLTI 586
 QY 589 TAADPEGLISYSLRDPTGMLKIDHVTGEIRSVALD-REAGSPRYOVVATVEVGGSSL 647
 DB 587 QPSDPISRTLRFLSLVNDSEGMCIERFSGEYHVAQSLQAQPEDITYLVEADTA---- 642
 QY 648 SSVSEPHLLIMDVNDPRLAIDYGLFPCHPILASPSLIF-EATDDQHLFRGPHFTFS 706
 DB 643 -----LTLAPVPSQ-----YICTRPDQGLVSGSKPDLASGHP-YSTF 683
 QY 707 LG-SGSLQNDWEVSKINGTHARLSTRHTDPEERAVVLLIRINDGSRPLEGIVSLPVTPC 765
 DB 684 LGPNPTVQDRDWRQLQTLNGSHA YLLTALHWVEPREHII PVVSHNAQ---MMQLLVAVIYC 740
 QY 766 SC-VEGSCTRPAGHQGTGTPVGAAGIILLTLLVIGIILAVPIRIKDKGKDNVESQAQ 824
 DB 741 RCNVEGQCMRKVGRMRKMPKLSAVGILVGTVAIGIFILIFTHWMSRKXDPDPAOS 800
 QY 825 SEYK 828
 DB 801 VPLK 804
 RESULT 6
 ID AAU83685
 ID AAU83685 standard; Protein; 807 AA.
 XX AAU83685;
 AC 08-MAY-2002 (first entry)
 DT XX
 DE Human PRO protein, Seq ID No 188.
 XX
 KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
 KW breast cancer; prostate tumour; rectal tumour; liver tumour;
 KW pericyte cell proliferation; chondrocyte cell proliferation;
 KW tumour necrosis factor-alpha.
 OS Homo sapiens.
 XX
 XX WO200208288-A2.
 PD 31-JAN-2002.
 XX
 PF 29-JUN-2001; 2001WO-US21066.
 XX
 XX 20-JUL-2000; 2000US-219556P.
 PR 25-JUL-2000; 2000US-220585P.
 PR 25-JUL-2000; 2000US-220605P.
 PR 25-JUL-2000; 2000US-220607P.
 PR 25-JUL-2000; 2000US-220624P.
 PR 25-JUL-2000; 2000US-220638P.
 PR 25-JUL-2000; 2000US-220664P.
 PR 25-JUL-2000; 2000US-220666P.
 PR 26-JUL-2000; 2000US-220893P.
 PR 26-JUL-2000; 2000WO-US20710.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23528.
 PR 15-SEP-2000; 2000US-000000P.
 PR 10-NOV-2000; 2000WO-US30673.
 PR 28-NOV-2000; 2000US-253646P.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000US-0747259.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 10-MAY-2001; 2001US-0854280.

PR 25-MAY-2001; 2001WO-US17092.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 XX WPI, 2002-172001/22.
 DR N-PSDB; ABK33629.
 XX
 PT One hundred and twenty two nucleic acids encoding PRO polypeptides,
 PT useful for treating a PRO related disorder and for diagnosing tumors
 PT such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
 PT tumour or liver tumour -
 XX
 XX Claim 11; Figure 188; 359pp; English.
 XX
 CC The invention relates to one hundred and twenty two nucleic acids
 CC encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
 CC encode human secreted proteins. The PRO nucleic acids, polypeptides,
 CC agonists and antagonists are useful for treating a PRO related disorder.
 CC The PRO polypeptides are useful for diagnosing tumours, especially lung
 CC cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. The PRO polypeptides are useful for stimulating the
 CC proliferation of, or gene expression, in pericyte cells, for stimulating
 CC the proliferation or differentiation of chondrocyte cells, for
 CC stimulating the release of tumour necrosis factor-alpha from human blood,
 CC for stimulating or inhibiting the proliferation of normal human dermal
 CC fibroblast cells. The PRO polypeptide may also be used as molecular
 CC weight markers and for tissue typing. The PRO nucleic acids have
 CC applications in molecular biology, including use as hybridisation probes,
 CC and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
 CC protein sequences of the invention.
 XX
 XX Sequence 807 AA;
 SO
 Query Match 21.0%; Score 908.5; DB 23; Length 807;
 Best Local Similarity 30.2%; Pred. No. 1.1e-64;
 Matches 237; Conservative 159; Mismatches 335; Indels 53; Gaps 21;
 QY 62 LTGE---TDNIVIEIE-GLLYNRALDRETSNHLQVLAADANGIIEGPPITIEV 116
 DB 57 LSGDSQKATGEPAMPDGSFLVTRALDREEQLEYQLQVLEMDQDHVLMGPVLYVHV 116
 QY 117 KDINDNRPTLQSKYEGRVQNSRPGKFLYVNAATDDPATPNQOLYQIVQLPMINN 176
 DB 117 KDENDQVPHFSQAIYRKLRSRGRFGIFLEASPDDEGTANSDLRPHILSQAPQPS 176
 QY 177 VMTFQINNTGAI SLREGSQELNPAKNPSYNLVISYKMGQSGSENSFSDTSDVIATVE 236
 DB 177 PDMFOLEPRILGALALSPKSGTSLDHALERTYQLLVQKMGDDA-SGHQATATVEVSIIE 235
 QY 237 NIKAKRPVEMVENSIDPHRIKITQVRANDPGAQYSLVDKEKIPRPFPSIDQGDIIYVQ 236
 DB 236 STVWSLEPHILANLKVLYPHHMAQVMSGDDVHYHL---ESHPPGPFVNAEGNLYVTR 232
 QY 297 PLDREKDAVYFAVAKDEYKPLSYPLIHKVVKIINDNPPCPSPVYFEVQENRILG 356
 DB 293 ELDRBAQAEYLLQVRQNSHGEDIYAPLEHLVLMQNDNVPIICPPRPDVSJIPELSPPG 352
 QY 357 NSIGITLTAHDREDENTANSFLNRYVEOTPKLPMDG-LFLIQTVAQMLQAKSLKK-QD 414
 DB 353 TETVTRLSAEDADA PGSPNSHVYQLSPEDDEGEGAFQVDPITSGSVTLGVLPKRAQN 412
 QY 415 TPQYNTLIEVSDQ--FKTLCFQVQINVIDNDQIPFEKSDYGNLTAEPTNIGSTILTI 472
 DB 413 ILLLVAMDLAGEGSSSTCEVEAVATINDHAPEFTSQIPISLPEDEVEPGTLVAML 472
 QY 473 QATDAD-EPEFTGSSKIL-YHIIKGDSEGRGLGVDPDPTNNGYVLIK-KPLDEPETA VSN 528
 DB 473 TALDADLEF---AFRLMDPAIERGDTIGTGLDWEF--DSGHVRLKCKNLSTEAAPSH 527
 QY 529 IYFKAENPEELVEGVKNASSFAKFTLITVDVAEAPQSFQHVFOAKVSEDAIGTKVGNV 588

DB 528 VVVVVQSVAKIV-GPSPGPATATVTVLVERVMPPKLQDESYSYASVPIAPAGSFLTRI 586
 QY 589 TAKDPGLDISYSLRGDTGMLKIDHVTGEIFSAVLDD-REAGSPRYQVAVATEVGGSSL 647
 DB 587 QPSDPLSRTRFSLVNDSEGLCTEKPSGSEVHTAOSLQGAQPDITVTLVEADDTA---- 642
 QY 648 SSVSEPHLLIMDVNDNPPRLAKDYTGILFCFPLSAPGSLIF-BATDDQHLFGPPHFTFS 706
 DB 643 -----LTLAPVPSQ-----YLCTPRQDGLIVSGSPKDPDLASGHP-YSTFT 683
 QY 707 LG-SGLQNDWEVSKINGTHARLSTRHTDEEPAVYVLRINDGRRPLEGIVSLPYTFC 765
 DB 684 LGNPPTVQRDMRLQTLNGSHAYLTLALHWEPREHIIIPVVSHNAQ---MMQLLVAVIYC 740
 QY 766 SC-VESGCRPAGHQGIPTVGMAVGILLTLLVIGIILAVFIRIKDKGKNVESAQA 824
 DB 741 RCNVEQCKRKVKRGMKMPKLSAVGILVGLTVAIGIFLILITHTMSRKPDQPADS 800
 QY 825 SEVK 828
 DB 801 VPLK 804
 RESULT 7
 AAU09959 standard; Protein; 830 AA.
 ID AAU09959
 AC AAU09959;
 DT 18-JUN-2002 (first entry)
 DT XX
 DE Human cadherin-2 (CDHN-2) protein.
 XX
 KW Cadherin; CDHN-2; nootropic; neuroprotective; cardiant; human;
 KW antiinflammatory; gastric; immunostimulant; cytostatic; immunogen;
 KW Alzheimer's disease; Parkinson's disease; multiple sclerosis;
 KW atherosclerosis; angiogenesis; musculoskeletal; ataxia; myotonia;
 KW gastrointestinal; gastritis; insulin dependent diabetes mellitus;
 KW Crohn's disease; inflammatory; asthma; rheumatoid arthritis; lupus;
 KW cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide /note= "signal peptide"
 FT Protein /note= "Mature CDHN-2"
 FT Domain /note= "Mature CDHN-2"
 FT Domain /note= "Transmembrane domain"
 FT Domain /note= "Transmembrane domain"
 FT Domain /note= "Transmembrane domain"
 FT Domain /note= "Transmembrane domain"
 FT Domain /note= "Transmembrane domain"
 XX
 PN W0200179293-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 18-APR-2001; 2001WO-US12687.
 XX
 PR 18-APR-2000; 2000US-198466P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 DR WPI; 2001-611722/70.
 DR N-PSDB; AAS15602.
 PT Isolated cadherin polypeptide useful for the treatment of central
 PT nervous system, cardiovascular, musculoskeletal, gastrointestinal,

PT inflammatory, immune system and cell proliferation disorders -

XX Inflammatory, immune system and cell proliferation disorders -

PS Claim 13; Figure 8; 147bp; English.

XX This sequence represents the cadherin-2 (CDN-2) protein of the

CC invention. This invention comprises the nucleotide and protein

CC sequences of two members of the cadherin family of membrane

CC glycoproteins CDN-1 and CDN-2. Also included in the invention is

CC a method for producing the CDN proteins by culturing a transfected

CC host cell line with a vector containing the CDN proteins and an

CC antibody which selectively binds to the protein. The CDN proteins

CC may have immunotropic, neuroprotective, cardiant, antiinflammatory,

CC gastric, immunostimulant and cytostatic activity and can be used

CC as cadherin-modulators using antisense-therapy or gene-therapy.

CC These proteins are useful for identifying compounds which bind

CC or modulate CDN-1. The proteins, nucleic acids and antibodies

CC are useful for treating a subject with a disorder characterised

CC by aberrant or unwanted cadherin protein or nucleic acid activity.

CC These disorders include central nervous system (e.g. Alzheimer's

CC disease, Parkinson's disease, multiple sclerosis), cardiovascular,

CC (e.g. atherosclerosis, angiogenesis), musculoskeletal (e.g. ataxia,

CC myotonia) gastrointestinal (e.g. gastritis, insulin dependent

CC diabetes mellitus, Crohn's disease) (e.g. inflammatory and

CC immune system (e.g. asthma, rheumatoid arthritis, lupus) or cell

CC proliferation disorders (e.g. cancers, leukaemia) and many other

CC diseases/disorders listed in the specification.

XX

XX Sequence 830 AA;

Query Match 20.3%; Score 877.5; DB 22; Length 830;

Best Local Similarity 30.5%; Pred. No. 3.7e-62;

Matches 245; Conservative 141; Mismatches 353; Indels 63; Gaps 22;

QY 62 LYGSTD---NIFYERB-GLLYNRAIDRETSNMLQVAALDANGIYEGPPIITEV 116

DB 59 LSGDSNTADQNTAVDTDSGLVATRTLDREKKEAYQLQVLTSESDRILMGPOLVTVHV 118

QY 117 KDINDNPTPLQSKYEGSVQNSRPGKPELVVNTDLDPTPGQLYQYQVVIDLPINN 176

DB 119 KDENDQVPOPSQALYRAQLSGQTRPGVFLFLESDDGAPETANSDFRLLSSPOPL 178

QY 177 VMYFOINNKGAISLTREGSOELNPAKNPSYLVISVDMGSGSENSFSDTSVDIIVTE 236

DB 179 PDWQDLPFLGALALSPSGSTSLDHALERTYQLLVQYKMDQDSGHAQAT -VEISIVE 237

QY 237 NIWKAPEVEMENSTDPHPKITQVRNNDGQAYSLVDEKELFRFPESIDQEGDIYVTO 296

DB 238 NSWAPLEPVLAEMLKVYYPHSIAQVMSGGDVHYOL--ESQPPGPFDDVTGEMLVYTM 294

QY 297 PLDEEKDAYVFAVNDVEGKPLSYPLEIHVKVDINDNPTCPSPVTVREVOENRLG 356

DB 295 ELDEBAQAEYQLQVRAQNSHGEDYAEPLHVVVNDENNAVPSPHDPTVNIPELSPPG 354

QY 357 NSIGTLTAHDEDEMTANSFLNRYIVEQTPKLPMDG-LFLIQYAGMLQL-----AKOS 409

DB 355 TEIARLGAEDLDAGSPNSHIVYQLLSPEBEAGAKNAFELDPSSGSLGTAFRLHGS 414

QY 410 LKKQDTPQYNLTIEV--SDKFTLCFVQVNVINDDOIPIFEKSDYGNLTLLADTNIGS 467

DB 415 ILAQ-----VLAVDLAGESEGSLSTCEVTAVVTVNNHAPFINSOIGPVLTPEDVKG 469

QY 468 TILTIQATDAD-REPTGSKTL-VHIKDGSEGLGVDTDHTNTGVYIIKKPIDEFRA 525

DB 470 LVATLMTADLEP--AFRLMDPAIEBGDEGIFDLSWEDSDHVOURLKKNLSYEAP 526

QY 526 VSNIVFAENDEPVPFGVKNVASSFAKFTLIVTVNENAPQSCVHFQAKVSEDAIGTKV 585

DB 527 DHKVVVVSNIIELV-GRGPPAATATYTIIVENVVAPLKIDQSYERSIVSTPASL 585

QY 586 GNVTAQDEGLDISYLRGTRGMKIDHTVGEIFSVAPLD-REAGSPYRVQVVA--TEV 642

DB 586 LTIQPSDPMSTRILRFSLVNDSEGWLCKIEVGEVHTAQSLQGAQPGDTYTVLVAQDTDK 645

QY 643 GGSLSLSSVESEFHLI-----LMDVNDNPPRLANDYTGLEFFCHPLSAPGSLIFEATDDQH 696

DB 646 PGLSTASATVVIHFLKASPVPAITLSAGPSR-----HLCTPRQDYGVVSGVSDPD 697

QY 697 LFRGPHTTSLG-SGSLQNDMEVSKINGTHARLSTHTPEERAYVVLRI-NDGSRPL 754

DB 698 ANNGCYSPALGNPNPVQDWRLOPLNDSHAYITLALHWPEPEYVPPVVDHDTMMQL 757

QY 755 EGVSLPVPFCSC-VEGSGERPAGHOTGIPTVMAGVILITLLVIGIILAVVFIRI--- 810

DB 758 Q---VKAVVCRNVGQCRKVKRKMPTKLSAVGLLGLTALIGFILLVFTHLALA 813

QY 811 KKDGRDNVESAGASEVKPLRS 832

DB 814 RKLD-----QPADSVPLKA 828

RESULT 8

AAV70741

ID AAV70741 standard; protein; 906 AA.

XX AAV70741;

XX 24-JUL-2000 (first entry)

XX Human N-cadherin.

XX Human N-cadherin.

XX Human N-cadherin; Wnt antagonist; contraceptive; contraceptive vaccine;

XX oocyte development; female primate contraception; oocyte viability;

XX monoclonal antibody; Wnt signalling.

XX Homo sapiens.

XX

XX Key Location/Qualifiers

XX Domain 160..724

XX Domain /label= Extracellular_domain

XX Domain 747..906

XX /label= Cytoplasmic_domain

XX W0200021555-A1.

XX 20-APR-2000.

XX 13-OCT-1999; 99MO-US23640.

XX 15-OCT-1998; 98US-0104355.

XX (HARD) HARVARD COLLEGE.

XX McMahon AP, Parr BA, Valino S;

XX WPI; 2000-317845/27.

XX Contraceptive composition for inhibiting oocyte development in a female

XX primate comprises a Wnt polypeptide antagonist -

XX Example 3; Page 27-28; 57bp; English.

XX The patent discloses a method of female primate contraception comprising

XX administering an antagonist of a Wnt polypeptide, inhibiting oocyte

XX development. Wnt polypeptides are useful for promotive maturation of an

XX immature oocyte. Wnt polypeptides are also useful for increasing the

XX number of mature oocytes and to enhance oocyte viability. Soluble

XX fragments of Wnt polypeptides have the ability to inhibit Wnt signalling,

XX e.g., by blocking binding of a naturally-occurring Wnt protein to its

XX receptor. They may be used to generate monoclonal antibodies which can

XX inhibit oocyte development. The present sequence is the human N-cadherin.

XX Signal transduction by beta-catenin is mediated by binding to the

XX cytoplasmic domain of a cadherin. Dominant negative mutants of N-cadherin

XX which inhibit Wnt-1 class signals can also be used as Wnt antagonists.

XX Sequence 906 AA;

Query Match	16.2%	Score 699.5	DB 21	Length 906
Best Local Similarity	27.4%	Pred. No. 1.3e+47		
Matches 210	Conservative 122	Mismatches 331	Indels 103	Gaps 21
QY	117	KDINNRPFPLDSKXSGVRSRSGKPEPLYNATDLDPATPNQLYQYIQLPMINN	176	
DB	47	KDVHGQRP-LANVKF-----SNCKGRKKQYVSSSEPPADPKVDEGDMY--AVRSFPLSE	98	
QY	177	VMYFOINNKGTGAILSTREGSGELNPAKPSYNVLVISVYDMGGQSENFSDDTSVDIIV--	234	
DB	99	HAKFLIY-----AQQKEQTEQKQVAVKLSLKP--LTESVYKSAEVEIYFP	144	
QY	235	-----TENIWKAPKEVENENSTDPHPIKTIQVQRN-----DPGAQ	270	
DB	145	ROFSKRSGHLOKQKDDWIP-PINLPENSRGFPPEOLVIRISDRDKNLSRYSVTGPA-	202	
QY	271	YSLVMEKELPREPFSDQ-EGDIYTOPLDREKAAVYFAVAKDEYKPLSYPLEIHYK	329	
DB	203	-----DQPPGIFLINPISGQISTYTKPLDRQQLRPHLRHAVIDINGNOVENPIDIYIN	256	
QY	330	VKDINDNPTCPSPYTFVEVQGENEISAGSITGLTHADREDEENTANSFLNRYIVEQPKLP	389	
DB	257	VIDKMDNRPREPLHQWNGTVPEGSRGFTVMTVTALIDADDPALNGLMRYIVSGAPRP	316	
QY	330	MDGLFLIQTYAGMLQAKOSLKKQDTPQYNTLIEVSDK-----FTLCPQIVNIDIN	443	
DB	317	SPNMFTTINNETDITTVAAAGLDREKVOQYTLIIQATDMEGNETYGLSNTATAVITVTDN	376	
QY	444	DOIPIF-EKSDVGNLTABDNTIGSTIITLQATDDEPFQSGSKLYHIIGDSEGRIGV	502	
DB	377	DNPEFTAMTYGEV--PENRVDITVANLTVDKQDPHTPMAVNAVIRISGDDPTGRFAI	433	
QY	503	DTDPTNTGYVLIKKPLDPEETAASNIYFKAKENPEPLFVGYNKSNASEFAKFTLLVTDVNE	562	
DB	434	QTDPRNSNGLVTVVXEDIDFETRMFVLTYAAMQVPLAKGIGHRPQSTATVYSVYIDVNE	493	
QY	563	APQSGQVFOAKVSDVAIGTKVGNVTADEPGL--DISYSLRGDTGMLKIDHTVGEI	619	
DB	494	NPYFAPNPKIRIOEBGLHAGTMLTFPTADDPDPRVQONIRVTKLSDPANMLKIDFVNQOI	553	
QY	620	FSVAALDEA-----GSPRVQVVAIEVGGSSLSVSEPHILIMDVNDNPPRLADYTLGF	675	
DB	554	TTIAVLADRESPVNKNINATFPLASDNGIPMPSGTGLQIYLLDINDAPVL-----	606	
QY	676	FCHPLSA-----PGLIFEATDDQHLFRGPH-FTFSLSGSLSDNMEVSKINGTAR	727	
DB	607	---PQEAETCEPNDNSINITRLDVIDIPNACFPAPFDLPSEVITIKRWITTRRLNGDPAQ	663	
QY	728	LSTRHTDEERAYVVLIRINDGRPDLGLVSLPYTFQSC-VEGSCFPAPAGHOTGIPTV-	785	
DB	664	INLIKTFLEAGIYEVPIITTDGPNPKSNISILRYKVQCQCSNDGC-----IDVDRIV	716	
QY	786	--GNAVGILLTLTLVYGIITLAVVFRIRKKDKKDNVESAQASEVPR	829	
DB	717	GAGLGATGAIIALICIIILILIVLMFVMMKRRDKEROAKOLLIDP	762	
RESULT 9				
ID	ABG14316			
XX	ABG14316 standard; Protein; 906 AA.			
XX	ABG14316;			
XX	AC			
XX	DT			
XX	18-FEB-2002 (first entry)			
XX	Novel human diagnostic protein #14307.			
XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;			
XX	food supplement; medical imaging; diagnostic; genetic disorder.			
XX	Homo sapiens.			
XX	W0200175067-A2.			

[illegible]

QY 444 DOIPIF-EKSDYGNLTLAEDNTIGSTLITQATDADPEFTGSSKILYHIKGDSEGLGV 502
 DB 377 DNPEFTAMTFYGEV---PENRVDIIVANLTVTDKQDPHPANNAVIRISGDPFGFAI 433
 QY 503 DTDPHNTGYVYIIKKPLDFETAAVSNIVFKAENPEPLVFGVKYNASSFAKFTLIVTDVNE 562
 DB 434 QTDPSNDGLVTVVKPIIDFETNMFVLTVAAENQVPLAKGIQHPPOSTATVSVTVIDVNE 493
 QY 563 APOFSQHVFOAKSSEDAIGTKVGNVTAADPEGL---DISYSLRGDTRGMKIDHTVGEI 619
 DB 494 NPYFAPNPVKIIRQEEGHAGTMTLTTFAODPDRMOONIRYTKISDPANMLKIDPVNGOI 553
 QY 620 FSVAPLDREA---GSPYRQOVVATEVGGSSLSVSEPHILMDVNDNPPRLAKDYTGLF 675
 DB 554 TTIAVLDRBSPNKNINYNATFLASDNGIPRMSGTGLQIYLDINDMAQVLT----- 606
 QY 676 FCHPLSA-----PGSLIFEATDDQHLFRGPH-FTFSIGSSGLQNDWEVSKINGTHAR 727
 DB 607 ---POEATCETPDPSNINITALDYDIDPNAGPFAFDLPSPVITIKENWITTRLNGDPAQ 663
 QY 728 LSTRHDPFERAVVLRINDGRRPPLGIVSLPVTFCSG-VGSSCFRPAHQGTGIPTV- 785
 DB 664 LNLKIFLEAGIYEVPIITDTSNPKPSNISILRVKVCQCCDSNGDC-----TDVDRIV 716
 QY 786 --GMAVGILTLTLVIGLILAVFIRIKKQKQDNVESAQSEVKP 829
 DB 717 GAGLGTAITAILCTIITLITLILVFMVMMKRKRKROKOLIDP 762
 RESULT 10
 ABB81474
 ID ABB81474 standard; Protein; 906 AA.
 AC ABB81474;
 XX 30-AUG-2002 (first entry)
 DT Human N-cadherin protein SEQ ID NO:46.
 XX Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;
 KM o-catenin; colon cancer.
 XX Homo sapiens.
 OS
 XX US2002045591-A1.
 PN 18-APR-2002.
 PD 17-JUL-2001; 2001US-0905963.
 PF 26-MAY-1998; 98IL-0124650.
 PR 26-MAY-1999; 99US-0318633.
 XX (GEIG/) GEIGER B.
 PA (BENZ/) BEN-ZE'EV A.
 PA (SADO/) SADO E.
 XX Geiger B, Ben-Ze'ev A, Sado E;
 PI WPI; 2002-499105/53.
 DR N-PSDB; ABN89391.
 XX
 PT New construct encoding soluble cytoplasmic portion of cadherin
 PT including beta catenin binding domain useful in treating cancer
 PT associated with high beta-catenin activity e.g. colon cancer and
 PT melanoma
 XX
 PS Example 3; Page 45-48; 102pp; English.
 CC The present invention describes a pharmaceutical composition for treating
 CC cancer associated with abnormally high beta-catenin activity. The
 CC pharmaceutical composition comprises a gene therapy vehicle harbouring a

CC polynucleotide that contains: (1) a nucleotide sequence encoding a
 CC soluble cytoplasmic portion of a cadherin which lacks a transmembrane
 CC portion and an extracellular portion of the cadherin, and includes a
 CC beta-catenin binding domain; and (b) an upstream promoter for directing
 CC expression of the soluble cytoplasmic portion of the cadherin in a
 CC mammalian cell. Also described is a pharmaceutical composition for
 CC treating cancer associated with abnormally high activity levels of
 CC beta-catenin comprising a gene therapy vehicle harbouring a
 CC polynucleotide that contains: (a) a nucleotide sequence encoding an
 CC o-catenin; and (b) an upstream promoter for directing expression of the
 CC o-catenin in a mammalian cell. The pharmaceutical compositions have
 CC cytostatic activity and can be used in the suppression of
 CC beta-catenin-mediated transactivation. They can be used for treating
 CC cancers associated with abnormally high activity levels of beta-catenin
 CC such as colon cancers and melanomas, by reducing these high activity
 CC levels of beta-catenin in mammalian cells. The present sequence
 CC represents human N-cadherin which is used in the exemplification of the
 CC present invention.

Sequence 906 AA;

Query Match 16.2%; Score 699.5; DB 23; Length 906;
 Best Local Similarity 27.4%; Pred. No. 1,3e-47;
 Matches 210; Conservative 122; Mismatches 331; Indels 103; Gaps 21;

QY 117 KDINDRPFPLQSGKESVYKQNSRPGKFLVYVATDLDPATPNGQLYQIVQLPMINN 176
 DB 47 KDVBGQP-LLNKF-----SNCGKRKYQYESSPADPVDDGVV--AVNSFPLSSE 98
 QY 177 VMYFOINNTGALSTRSGQELINPAKNPSYLVISVKMGSGSENSPDDTSDVILV-- 234
 DB 99 HAKFLIY-----AQDKETQEKQVAVKLSLKP--LTSESVSESAEEVERIVFP 144
 QY 235 -----TENWKAAPKPYEMTENSDDPHKITQVRN-----DPGAQ 270
 DB 145 RQFSKSHGLQKQKRWVIP-PTNLPENSGRPPQELVTRSDRDKNLSIRYVTGEGA- 202
 QY 271 YSLVDKEKLPFPFSIDQ-EGDIYVTOPLDREKDAVVFYAVAKDEYKPLSYLEIHVK 329
 DB 203 -----DQPPGTGFIINPISGQSVTKPLDRBQIARFLRAHVDINGQVENPIDIVIN 256
 QY 330 VKDINDNPPTCSPSPVTFVEQENERLGNISGTYLTAHDBRENTANSFLNRYIEQTPKL 389
 DB 257 VIDMINDRPEFLHQVWNGTVPEBSKPGTYVMVTALIDADPNALMGRLRYRIVSOABSTP 316
 QY 390 MDGLFLLQTYAGMLQAKQSLKKQDTPQYNLTIEVSKD-----FKTLCFVQINVYDIN 443
 DB 317 SPNMFTINNETGDIITVAAGLDREKVOQYTLITQATDMESIPYGLSNTATAVIYTVDN 376
 QY 444 DOIPIF-EKSDYGNLTLAEDNTIGSTLITQATDADPEFTGSSKILYHIKGDSEGLGV 502
 DB 377 DNPEFTAMTFYGEV---PENRVDIIVANLTVTDKQDPHPANNAVIRISGDPFGFAI 433
 QY 503 DTDPHNTGYVYIIKKPLDFETAAVSNIVFKAENPEPLVFGVKYNASSFAKFTLIVTDVNE 562
 DB 434 QTDPSNDGLVTVVKPIIDFETNMFVLTVAAENQVPLAKGIQHPPOSTATVSVTVIDVNE 493
 QY 563 APOFSQHVFOAKSSEDAIGTKVGNVTAADPEGL---DISYSLRGDTRGMKIDHTVGEI 619
 DB 494 NPYFAPNPVKIIRQEEGHAGTMTLTTFAODPDRMOONIRYTKISDPANMLKIDPVNGOI 553
 QY 620 FSVAPLDREA---GSPYRQOVVATEVGGSSLSVSEPHILMDVNDNPPRLAKDYTGLF 675
 DB 554 TTIAVLDRBSPNKNINYNATFLASDNGIPRMSGTGLQIYLDINDMAQVLT----- 606
 QY 676 FCHPLSA-----PGSLIFEATDDQHLFRGPH-FTFSIGSSGLQNDWEVSKINGTHAR 727
 DB 607 ---POEATCETPDPSNINITALDYDIDPNAGPFAFDLPSPVITIKENWITTRLNGDPAQ 663
 QY 728 LSTRHDPFERAVVLRINDGRRPPLGIVSLPVTFCSG-VGSSCFRPAHQGTGIPTV- 785
 DB 664 LNLKIFLEAGIYEVPIITDTSNPKPSNISILRVKVCQCCDSNGDC-----TDVDRIV 716

QY 786 --GNAVGLITLTVIGIILAAVFIKDKDKDNVESQAQSEVKP 829
 DB 717 GAGLGTGAIIAIIILCIIILILVLMFVMMKRRDKERQAKQLLIDP 762

RESULT 11
 ID ABB81471 standard; Protein: 912 AA.
 AC ABB81471;
 DT 30-AUG-2002 (first entry)

XX 30-AUG-2002 (first entry)
 DE Chicken N-cadherin protein SEQ ID NO:2.

XX Cadherin; alpha-catenin; cancer; beta-catenin binding domain; melanoma;
 KM o-catenin; colon cancer.

OS Gallus gallus.

PN US2002045591-A1.

PD 18-APR-2002.

PF 17-JUL-2001; 2001US-0905983.

PR 26-MAY-1998; 98IL-0124650.

PR 26-MAY-1999; 99US-0318633.

PA (GEIGI) GEIGER B.
 (BENZ) BEN-ZE'EV A.

PA (SADO) SADO E.

PI Geiger B, Ben-Ze'ev A, Sado E;

DR WPI, 2002-499105/53.

DR N-PSDB; ABBN9350, ABBN9351.

PT New construct encoding soluble cytoplasmic portion of cadherin
 including beta catenin binding domain useful in treating cancer
 PT associated with high beta-catenin activity e.g. colon cancer and
 PT melanoma

PS Example 1; Page 19-21; 102pp; English.

XX The present invention describes a pharmaceutical composition for treating
 CC cancer associated with abnormally high beta-catenin activity. The
 CC pharmaceutical composition comprises a gene therapy vehicle harbouring a
 CC polynucleotide that contains: (1) a nucleotide sequence encoding a
 CC soluble cytoplasmic portion of a cadherin which lacks a transmembrane
 CC portion and an extracellular portion of the cadherin, and includes a
 CC beta-catenin binding domain; and (b) an upstream promoter for directing
 CC expression of the soluble cytoplasmic portion of the cadherin in a
 CC mammalian cell. Also described is a pharmaceutical composition for
 CC treating cancer associated with abnormally high activity levels of
 CC beta-catenin comprising a gene therapy vehicle harbouring a
 CC polynucleotide that contains: (a) a nucleotide sequence encoding an
 CC o-catenin; and (b) an upstream promoter for directing expression of the
 CC o-catenin in a mammalian cell. The pharmaceutical compositions have
 CC cytostatic activity and can be used in the suppression of
 CC beta-catenin-mediated transactivation. They can be used for treating
 CC cancers associated with abnormally high activity levels of beta-catenin
 CC such as colon cancers and melanomas, by reducing these high activity
 CC levels of beta-catenin in mammalian cells. The present sequence
 CC represents chicken N-cadherin which is used in the exemplification of
 CC the present invention.

XX Sequence 912 AA;

Query Match 16.1%; Score 696; DB 23; Length 912;
 Best Local Similarity 26.9%; Pred. No. 2.6e-47;
 Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

QY 126 FLOSKTEGVSRONSREKPEFLYYNATLDDPATPNQOLYYQIVIOLEPMINWYFOIN-- 183
 DB 41 FPDVSAVAVSRSGVGGQFLNVRFGSCDE-----NRKIFGSGSEP 81

QY 184 ----NKGAI SLTREGSOELNPAKPSYLVLSVDMGQS-----ENSPSDDT 228

DB 82 EDRFVGEDGVVAERSFQLSAEPTL-----FVVSADKETOEBOMKVKLTPPAFAFGAS 136

QY 229 SVDIIVTENT-----WKAKPPEVMVENSSTDPHPKITYRWMDPGA 269

DB 137 EKQKKIEDIIFWQYKQSSHLKQKRDVIVP-PINLPENSGPFPQELVIR----- 189

QY 270 QVSLVDKELRPPPSI-----DQ-----EGDIYVTOPLDREKDAVFAVAK 313

DB 190 ----SDRKSLSLRYSVTGAGADQPPGIFINIPISGQSLVTPKDLREQLASPHLRHAHV 245

QY 314 DEYKPLSYPLLEHVKKYKIDNDNPPCPSPSYVFEVQENRCLNSIGTLTAHDDDEBNTA 373

DB 246 DVNGNQVENPIDIVINVIDMNDNRPEFLHQVWNGTPEGSKPQTYVWYTAIADADDPNAQ 305

QY 374 NSFANRYIVEQTEKLMMDGLFLQTYAGMLQAKOSLKKQDTPQYNULTIEVSKD----- 428

DB 306 NGMLRYRILSQABSSPSNNFTINNETGDIITVAAGLDREKVQOYTLITQATMEGNPTY 365

QY 429 -FKTLGFVQINVIDINDQIPIF-EKSDYGNLTAEEDTNGSTIITQATDADPEPTGSSK 486

DB 366 GLSNTADAVITVDVNDNPEFLPAMTFYGSV---PENRDVIVANLTVTDKDDPHPPAMN 422

QY 487 ILYHIITKDSSEGLGVDTPDHTNTGYIILKKPLDFTFAVSNIVFAENPEPLVFGVKIN 546

DB 423 ARQMTGGDPTGQFTILTPDNSNDGLVTVKRPIDFETINMFVLVAENQVPLAKGIQHP 482

QY 547 ASSFAKFTIIVTDVNEAPQSOHVFOAKVSEDAVIGTKVGNVAKXP-----EGLDISYSL 602

DB 483 POSTATVSVITVDVNSPFLVPPPKLVROBEGGLAASMLTTFARDPRMOQTSILYSK 542

QY 603 RGDTRGMLKIDHYTGEIFSVAPLDREA---GSPYVQVATEVGGSSLSVSEPHILIM 658

DB 543 LSPDPAWMLKIDPNNGIITTTAVLDRESIYQNNMVAATPLASNGIIPPMGCTTLOIYLL 602

QY 659 DVNDNPPRL-ANDYTLFFCHPLSAPGSLIFEATDDOHLFRGPHFTFSISGG--SLQND 715

DB 603 DINDNAPQVNPKEAT--TCETLQ-DNAINITAVDDPIDPNAGP-PAFELPDSPPSISKRN 657

QY 716 WEVSKINGHARLSTHTDPEERAYVYLRLINDGSRPPLEGIVSLVPTFSC--VEBSCFR 774

DB 658 WTVIRISGHAQLSLRPLRLEAGIDVPIVITDSGNPHASSTSVLKVXCQCDINDGC-- 715

QY 775 PAGHQTGPIPV--GNAVGLITLTVIGIILAAVFIKDKDKDNVESQAQSEVKP 829

DB 716 ----TDVDRIVAGLGTGAIIAIIILCIIILILVLMFVMMKRRDKERQAKQLLIDP 768

RESULT 12

ID ABB57233
 ID ABB57233 standard; Protein: 906 AA.

AC ABB57233;
 XX

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related protein sequence SEQ ID NO:606.

XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease.

OS Mus musculus.

PN WO200188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

```

XX 18-MAY-2000; 2000JP-0145977.
XX (UNYI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX Iahikawa K, Aseai S, Takahashi Y, Nagata T, Ishii Y;
XX WPI; 2002-034733/04.
XX N-PSDB; AB199593.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
XX expression levels of particular genes defined in the specification or
XX by determining the expression profile of a gene group comprising these
XX genes -
XX
XX Claim 2; Page 1560-1564; 2690PP; English.
XX
XX The present invention describes a method for examining ischaemic
XX conditions, comprising measuring the expression levels of particular
XX genes (I) in a test sample or determining the expression profile of a
XX gene group in the sample comprising genes selected from (I). The method
XX is useful for examining the ischaemic condition (e.g. compressive
XX ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
XX expression levels of particular genes (AB199202 to AB199912, encoding
XX the protein sequences in AB57020 to AB57374) or by determining the
XX expression profile of a gene group comprising these genes. The
XX expression levels or expression profiles produced by these genes are
XX used as an indicator when screening for ischaemic condition-improving
XX drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
XX represent PCR primers for a mouse ischaemic condition related sequence,
XX which are used in the exemplification of the present invention.
XX
XX Sequence 906 AA;
SQ
Query Match 16.1%; Score 694.5; DB 23; Length 906;
Best Local Similarity 27.6%; Pred. No. 3.4e-47;
Matches 211; Conservative 122; Mismatches 332; Indels 99; Gaps 23;
QY 117 KDINDNRPTLQSKESVSQNSRPGKPLVYNAITDDPATPGQLYQVIOLOPMINN 176
DB 47 KDVEGGP-LLNVFNSCNKR-----RKVOSESSEPADFKVDEGTVY--AVRSFPLTAE 98
QY 177 VMVFOI---NNKTG-----AISTREGSOELNPAKNP-SYVLVFSVDM--GQSENSF 224
DB 99 QAKLTYAQDKETOEKQVAVNLSREPTLTPEPKHEHIEIYFPQOLAKHSGALROK 158
QY 225 SDTTSVDIIVTENIWKAPKPEVWENSTDPHPKITOVN-----DPGAQYS 272
DB 159 RD-----WVIP-PINLPENSRGPPQELVIRSDRDKNLSLYSVTGPGA--- 202
QY 273 LVNKEKLPFPFSTIDQ-EGDIYVTOPLDREKDAVVFYAAKDYKGPLSELEIHKVK 331
DB 203 ----DQPTGFIILNPIISGQLSVTKPLDRELIAFPHLRAHAVDINGQVENPIDIVINVI 258.
QY 332 DINDNPTCPSPVTVFEVOENRNLGNSIGTLTAHDREBENTANSFLNRYIEQTPKIPMD 391
DB 259 DANDNRBEPLHQVWNGSVPESSKPGTYMTVTAIDADDPAALNMLKTRILISQAPSPSP 318
QY 392 GLFLIQTAYAGMLQAKOSLKKQDTPQYNLTIESDKD-----FKTLCQVQINVIDINDQ 445
DB 319 NMFITNETGDIITVAAGLDREKVOQVTLTIQATDMEGNPTYGSLNATAVITVTDVNDN 378
QY 446 IPIE-ESVDGNLTIAEDTNGITLTIQATDADBEPTGSSKILYHIKGSERLQVDT 504
DB 379 PPEFTAMTFYGEV---PENRVDVIVANILTVTDKQPHIPANNAVYRISGGOPTREFAILT 435
QY 505 DPHNTGTVTIKKPLDEFETAIVSNIVFEKAENPEPLVEGVKNAASFPKFTLIVDVNEAP 564
DB 436 DPNNDGLVTVYKRIIDETNRMFVLTAAENQVFLAKGIQRPPOSTATVSVTIVDVNEP 495
QY 565 QFSCHVQAKVSEDAVIGTKVGNVTAQPEGL---DISYSLRGDTRGMLKIDHTVGEIFS 621
DB 496 YFAFNPKTIRQEBGLHAGTMTLTTLTAQDPDRYMOONIRYTLSDPAWMLKIDPVNGQIT 555

```

```

QY 622 VAPLDREA-----GSPRYGVAVTEVGSSLSVSEFHLIMDVNDNPPRLAKDYTLGFC 677
DB 556 IAVLDRESRYVQNNITNATFLASDNGIPMSGCTGIQIYHLDINDNAPQVL----- 606
QY 678 HPLSA-----PGSLIFPATDDQHLFRGPH-FTFSSGSLQNDWEVSKINGTARLS 729
DB 607 -PQEAFTCETPEBNSINIALVDYIDPNAGPFPAFDLPSPVTKRWMTINRLNGDEPAQLN 665
QY 730 TRHTDPEERAYVYLITINNGRPPELEGYSLPYTFPSC-VEGSCFRPAGHGTGIPV--- 785
DB 666 LKIKFLBAGIYEPHITITDSGNPKSISILRYKVCQCSNBDP-----TDVRIVGA 718
QY 786 GMAVGIILFTLIVIGIILAVFPIRIKDKGKDNVESQAQSEVAP 829
DB 719 GLCTGAIIAILLCITLILIVLMFVWMKRRDKERQAKQLLIDP 762

RESULT 13
AAW25658
ID AAW25658 standard; Protein; 916 AA.
AC AAW25658;
XX 04-NOV-1997 (first entry)
DT Human cadherin-4.
XX Human; cadherin; rat; calcium-dependent cell adhesion protein;
XX superfamily; cytoskeleton; eatenin; cancer.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Msc-difference 78 /note= "Encoded by AAA"
FT
XX US5646250-A.
PD 08-JUL-1997.
XX 17-APR-1992; 92US-0872643.
XX 19-APR-1993; 93US-0049460.
XX 17-APR-1992; 92US-0872643.
XX 01-NOV-1994; 94US-0332638.
XX (DOHE-) DOHENY EYE INST.
PI Suzuki S;
XX WPI; 1997-362997/33.
XX N-PSDB; AAT85433.
XX
XX Human and rat cadherin polypeptide(s) - mediate cell-cell adhesion
XX
XX Example 2; Column 61-66; 56pp; English.
XX
XX This sequence represents human cadherin-4. The invention specifically
XX provides details of human cadherin-5, -8, -11, -12 and -13, and rat
XX cadherin-8, -5, -11 and -13. Cadherins are calcium-dependent cell
XX adhesion proteins. They are glycosylated integral membrane proteins
XX that have an N-terminal extracellular domain that determines binding
XX specificity, a hydrophobic membrane spanning region and a C-terminal
XX cytoplasmic domain, which is highly conserved among members of the
XX superfamily. The C-terminal domain interacts with the cytoskeleton
XX through eatenin and other cytoskeleton-associated proteins. The
XX novel cadherin proteins may be used in the analysis of the role of
XX cadherins in various cancers. Sequence analysis of the cadherin
XX proteins also allows investigation of the structure and function of
XX cadherin. The cadherin proteins may be isolated by using anti-cadherin
XX antibodies. These antibodies may also be used to modulate the activity
XX of cadherin and to determine the tissue specific distribution of cadherin

```


CC proteins. Each subclases of cadherins has a unique tissue distribution
 CC pattern.

XX Sequence 916 AA;

Query Match 15.8%; Score 684.5; DB 18; Length 916;
 Best Local Similarity 29.3%; Pred. No. 2.3e-46;
 Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKPEVWENSTDPHPKIKITQVRM---NDPCAQYSL--VDKELPRPFPSID-QEGDI 292
 171 WVIP-PINPENSNGPFPQOLVIRSDKNDIPRISITGVGADQPMVEVFSINSGRM 229
 293 YVTPQPLDREKDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPCTSPVTFVQEN 352
 230 YVTRPMDREHASYHLRAHAVDMNGKVENPIDLYIVIDMNDHPFINQVNCVDEG 289
 353 ERLGNSIGTLTAHDREENTANSFLNRYIVEQTPKLPMDGLFLIQYAGMLQAKOSLKK 412
 290 SKPTGYMTITANDADSTTANGMVRVRYIVTQPSQSMFTINSEBTGDIVTVAAGMDR 349
 413 QDPFOYNLTIEVSDKQ-----FKTLCFVOINVINDIQPIPEKSDYGNLTLAEDTNG 466
 350 EKVOQYTVIVQATDMEGNLYGSLNTATITVTDVNDNPFSEFTAFAG--EVPENSVE 407
 467 STLTITQATDADEPFTGSSKILYHIKDSBGRGVDTPHTNTGYVILKKPLDFETAAV 526
 408 TVVANLTVMRDQPHSNNMAVVRILISGDSGHSVATDPTNEGMTVVKADYELNRA 467
 527 SNIVFKAENPEPLVFGVKYKNAASFAPKFTLIVTDVNEAPQSOHVPOAKSEDAVIGTKVG 586
 468 FMLTVWVNSQAPLASGIQMSFOSTAGVTISIMDINEAPYPSNHKILRLEGVDPGTVL 527
 587 NTAKDPEGL---DISYSLRGDTRGMLKIDHVTGEISVAPLDREA---GSPYRQVVA 639
 528 TFSAVDDPRMOQAVRYSKLSDPASWHLINATNGQITVAVLDRESLYTNVYATFELA 587
 640 TEVGSGLSSVSEFHLILMDVNDNPPRLADYTGJFCHPLASAGSLIFEATDDOHLFR 659
 588 ADNGIPPASGTLQIYLIDINDAPELLKPAQICGRPLNLA---INITAADVHPNI 644
 700 GPH-FTFSLGSGSLQNDMEVSKINGTHARLSTRHTDFEERAVVYLIRINDGRRPLEGIV 758
 645 GPYVFELPVPVAPVARKWMTITRLNGDYAQLSLRILYLEAGMTDVPITVDSGNPLSNTS 704
 759 SLVPTFCSCVEGSCFRPAGHOTGIPVGMVAGILTLTLVIGIILAV--FIRIKDKK 816
 705 IIKVKVCPDNDGDCCTTIG--AVAAAGLGTGAIVAILICILITLTVLFLVMMKRREK 761
 QY 817 D 817
 Db 762 E 762

RESULT 14
 AAM13129
 ID AAM13129 standard; Protein; 916 AA.

XX AAM13129;
 XX
 DT 13-MAY-1997 (first entry)
 XX Full length human cadherin-4.

XX Ca2+ dependent; cell adhesion protein; foetal; cadherin; rat;
 KW brain; human; antibody; purification; determination;
 XX tissue expression; binding antagonist; calcium ion.

OS Homo sapiens.
 XX
 PN US559725-A.
 XX
 PD 28-JAN-1997.

XX 17-APR-1992; 92US-0872643.
 PF
 XX 26-JAN-1994; 94US-0188228.
 PR 17-APR-1992; 92US-0872643.
 PR 19-APR-1993; 93US-0049460.
 XX
 PA (DOHE-) DOHENY EYE INST.
 XX
 PI Suzuki S;
 XX
 DR WPI, 1997-108328/10.
 DR N-PSDB; AAT61920.
 XX
 PT Antibodies to cadherin proteins - useful as cadherin antagonists,
 etc.
 XX
 PS Claim 1; Columns 67-72; 59pp; English.
 CC The present sequence is full length human cadherin-4, which
 CC is a Ca2+ dependent cell adhesion protein. The human cadherin cDNA
 CC was isolated from a foetal brain cDNA library, using probes based
 CC on homologous rat cadherin cDNA.
 CC Antibodies or fragments that specifically bind the human cadherin
 CC can be used to purify the cadherin, determine its tissue expression
 CC and antagonise its ligand/antiligand binding activities.
 CC
 SQ Sequence 916 AA;

Query Match 15.8%; Score 684.5; DB 18; Length 916;
 Best Local Similarity 29.3%; Pred. No. 2.3e-46;
 Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKPEVWENSTDPHPKIKITQVRM---NDPCAQYSL--VDKELPRPFPSID-QEGDI 292
 171 WVIP-PINPENSNGPFPQOLVIRSDKNDIPRISITGVGADQPMVEVFSINSGRM 229
 293 YVTPQPLDREKDAYVFAVAKDEYKPLSYPLEIHVKVDINDNPCTSPVTFVQEN 352
 230 YVTRPMDREHASYHLRAHAVDMNGKVENPIDLYIVIDMNDHPFINQVNCVDEG 289
 353 ERLGNSIGTLTAHDREENTANSFLNRYIVEQTPKLPMDGLFLIQYAGMLQAKOSLKK 412
 290 SKPTGYMTITANDADSTTANGMVRVRYIVTQPSQSMFTINSEBTGDIVTVAAGMDR 349
 413 QDPFOYNLTIEVSDKQ-----FKTLCFVOINVINDIQPIPEKSDYGNLTLAEDTNG 466
 350 EKVOQYTVIVQATDMEGNLYGSLNTATITVTDVNDNPFSEFTAFAG--EVPENSVE 407
 467 STLTITQATDADEPFTGSSKILYHIKDSBGRGVDTPHTNTGYVILKKPLDFETAAV 526
 408 TVVANLTVMRDQPHSNNMAVVRILISGDSGHSVATDPTNEGMTVVKADYELNRA 467
 527 SNIVFKAENPEPLVFGVKYKNAASFAPKFTLIVTDVNEAPQSOHVPOAKSEDAVIGTKVG 586
 468 FMLTVWVNSQAPLASGIQMSFOSTAGVTISIMDINEAPYPSNHKILRLEGVDPGTVL 527
 587 NTAKDPEGL---DISYSLRGDTRGMLKIDHVTGEISVAPLDREA---GSPYRQVVA 639
 528 TFSAVDDPRMOQAVRYSKLSDPASWHLINATNGQITVAVLDRESLYTNVYATFELA 587
 640 TEVGSGLSSVSEFHLILMDVNDNPPRLADYTGJFCHPLASAGSLIFEATDDOHLFR 659
 588 ADNGIPPASGTLQIYLIDINDAPELLKPAQICGRPLNLA---INITAADVHPNI 644
 700 GPH-FTFSLGSGSLQNDMEVSKINGTHARLSTRHTDFEERAVVYLIRINDGRRPLEGIV 758
 645 GPYVFELPVPVAPVARKWMTITRLNGDYAQLSLRILYLEAGMTDVPITVDSGNPLSNTS 704
 759 SLVPTFCSCVEGSCFRPAGHOTGIPVGMVAGILTLTLVIGIILAV--FIRIKDKK 816
 705 IIKVKVCPDNDGDCCTTIG--AVAAAGLGTGAIVAILICILITLTVLFLVMMKRREK 761

QY 817 D 817
Db 762 E 762

RESULT 15

AAR63533 standard; Protein; 896 AA.

XX AAR63533;

DT 07-JUN-1995 (first entry)

XX Human HT-1376 cell-derived Leukocyte Chemotactic Factor.

XX rHT-LCF; HT-1376 cell derived leukocyte chemotactic factor;

XX immunopotentiator; monocyte and macrophage migration factor;

XX wound healing.

OS Homo sapiens (cell line HT-1376).

XX Key Location/Qualifiers

FT Peptide 1..30

FT Protein /label= signal_peptide

FT Protein /label= rHT-LCF

XX W09421809-A.

XX 29-SEP-1994.

XX 11-MAR-1994; 94WO-JP00397.

XX 15-MAR-1993; 93JP-0054349.

XX (CYTO-) INST CYTOSIGNAL RES INC.

XX Kawamura K, Watanabe K;

XX WPI; 1994-317029/39.

XX N-PSDB; AAQ72597.

XX Claim 5; Page 34-41; 68pp; Japanese.

XX Sequences coding for proteins having monocyte and macrophage

XX migration factor activity were isolated from a cDNA library

XX prepared from human HT-1376 cells (ATCC CRU-1472). The proteins

XX encoded by the isolated cDNA clones are useful as

XX immunopotentiators for treatment of infections, immune

XX deficiency diseases and cancer. The proteins are also useful in

XX wound healing preparations and for treatment of skin diseases.

XX Sequence 896 AA;

XX Query Match 15.5%; Score 671.5; DB 15; Length 896;

XX Best Local Similarity 30.8%; Pred. No. 2.5e-45;

XX Matches 187; Conservative 106; Mismatches 263; Indels 51; Gaps 22;

XX

XX

XX

QY 409 SLKKDPTPOYNLTIEVSDKPF-----TLCEVOINVINDIQPIPEKSDYGNLTAE 461

Db 310 YLDREAVVDKXSLIMKQVDMGQPFGLIGTSTCI--ITVDSNDNAFTFQNAE--AFVE 365

QY 462 DTNIGSTIILTIQATDADEPFTGSSKILYHIKGDSEGLGVDTPPTNGYVILKKPLDF 521

Db 366 ENAFNVEILRIPIEDDLDINTANMRVNFITLKNENGHKISTDKETNGVLSVAKPLAN 425

QY 522 ETAAVSNIVFKAEPEPLVFGV-KYASSPAKFTLIYTVNENAPQSQHVFAKVEDVA 580

Db 426 EENRQVNLIEIGVNNBAFPARDIPRYVALNRALVTVAHRDLDEGEECTPAQYVRIRKENLA 485

QY 581 IGTKGVNTAKOPE--GLDISLSLGDTRGMKIDHVGELFVSVAPLDREGSP----Y 633

Db 486 VSKINGYKAYVDENNGNGLRYKKLHDPKGMITIDEIGSITTSKILDBREVTPEKLEY 545

QY 634 RYQVVADEVGSSLSVSSEFHLIIMDVNDNPPRLAKDYGLFECPLSAPGSLIFEATD 693

Db 546 NITVLAIDKDDRECTGTAVN--IEDVNDNPEILGEY--VVICRPMGYTDIL--AVDP 599

QY 694 DQHLFRGPHFTFSLGSGS--LQNDWEVSKINGTHARLS-TRHTDFEERAYVILIRINDG 750

Db 600 DEEVHGAAP-FYFSLPMTSPEISRLMSLTIKVNDTPARLSYQKAGFQE--YTIPIYKD-- 654

QY 751 RPLLEGIVSLPYTFGSCVBSGCPRPAGHQGIPVGMAY-GIILTLTVIIGIILAVFIR 809

Db 655 RAGQAATKLRVNLCECTHTPQCRATSRSTGYLGKRALIALLIGIALLSFVLLTVCGV 714

QY 810 IKKDKG 816

Db 715 FGATKGR 721

Search completed: June 22, 2003, 01:04:37

Job time : 66 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 22, 2003, 01:03:29 ; Search time 35 seconds
(without alignments)
699.425 Million cell updates/sec

Title: US-10-025-380-1081

Perfect score: 4321

Sequence: 1 MIQAHLSICLMLTYLATG.....DKGDNVESQAQSEVPELRS 832

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5A COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5B COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6A COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4321	100.0	832	1 US-08-431-560-1	Sequence 1, Appl1
2	4321	100.0	832	1 US-08-463-345-1	Sequence 1, Appl1
3	696	16.1	913	2 US-08-474-067-6	Sequence 6, Appl1
4	696	16.1	913	2 US-08-474-068A-6	Sequence 6, Appl1
5	696	16.1	913	2 US-08-472-481-5	Sequence 5, Appl1
6	684.5	15.8	916	1 US-08-188-228-48	Sequence 46, Appl1
7	684.5	15.8	916	1 US-08-332-643-42	Sequence 42, Appl1
8	684.5	15.8	916	1 US-08-332-638-48	Sequence 48, Appl1
9	657.5	15.2	713	1 US-08-188-228-62	Sequence 62, Appl1
10	657.5	15.2	713	1 US-08-332-643-56	Sequence 56, Appl1
11	657.5	15.2	713	1 US-08-332-638-62	Sequence 62, Appl1
12	650.5	15.1	555	2 US-08-453-702A-98	Sequence 98, Appl1
13	634	14.7	837	2 US-08-474-067-7	Sequence 7, Appl1
14	634	14.7	837	2 US-08-474-068A-7	Sequence 7, Appl1
15	634	14.7	837	2 US-08-472-481-6	Sequence 6, Appl1
16	633	14.6	556	1 US-07-998-003A-98	Sequence 98, Appl1
17	633	14.6	556	1 US-08-453-374B-98	Sequence 98, Appl1
18	633	14.6	556	1 US-08-453-695A-98	Sequence 98, Appl1
19	633	14.6	556	1 US-08-268-161A-98	Sequence 98, Appl1
20	633	14.6	556	4 US-09-099-639-98	Sequence 98, Appl1
21	630.5	14.6	712	2 US-08-474-067-2	Sequence 2, Appl1
22	630.5	14.6	712	2 US-08-474-067-5	Sequence 5, Appl1
23	630.5	14.6	712	2 US-08-474-068A-2	Sequence 2, Appl1
24	630.5	14.6	712	2 US-08-474-068A-5	Sequence 5, Appl1
25	630.5	14.6	712	2 US-08-472-481-2	Sequence 2, Appl1
26	630.5	14.6	717	2 US-08-474-067-4	Sequence 4, Appl1
27	630.5	14.6	717	2 US-08-474-068A-4	Sequence 4, Appl1

28	630.5	14.6	717	2 US-08-472-481-4	Sequence 4, Appl1
29	608	14.1	884	2 US-08-474-067-8	Sequence 8, Appl1
30	608	14.1	884	2 US-08-474-068A-8	Sequence 8, Appl1
31	608	14.1	884	2 US-08-472-481-7	Sequence 7, Appl1
32	593	13.7	878	1 US-08-337-919-2	Sequence 2, Appl1
33	593	13.7	878	4 US-08-732-429-2	Sequence 2, Appl1
34	593	13.7	878	4 US-09-798-267-3	Sequence 3, Appl1
35	593	13.7	878	4 US-09-798-267-2	Sequence 2, Appl1
36	593	13.7	878	4 US-09-798-267-3	Sequence 3, Appl1
37	592.5	13.7	117	4 US-09-187-859-25	Sequence 25, Appl1
38	574.5	13.3	796	2 US-08-738-349-2	Sequence 2, Appl1
39	560.5	13.0	796	2 US-08-738-349-4	Sequence 4, Appl1
40	559	12.9	822	2 US-08-474-067-9	Sequence 9, Appl1
41	559	12.9	822	2 US-08-474-068A-9	Sequence 9, Appl1
42	559	12.9	822	2 US-08-472-481-8	Sequence 8, Appl1
43	555	12.8	106	1 US-08-326-117B-6	Sequence 6, Appl1
44	555	12.8	106	3 US-08-982-129-6	Sequence 6, Appl1
45	554	12.8	106	4 US-09-178-176B-11	Sequence 11, Appl1

ALIGNMENTS

```
RESULT 1
US-08-431-560-1
; Sequence 1, Application US/08431560
; Patent No. 5620855
; GENERAL INFORMATION:
; APPLICANT: Anne H. Dantzig, et al.
; TITLE OF INVENTION: Mammalian Influx Peptide
; TITLE OF INVENTION: Transporter
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/431,560
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/013,462
; FILING DATE: 04-FEB-1993
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 832 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-431-560-1

Query Match 100.0%; Score 4321; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIQAHLSICLMLTYLATGEGKFGSGPLKMTFSIYEGQSPQITTFQKMPAVTF 60
Db 1 MIQAHLSICLMLTYLATGEGKFGSGPLKMTFSIYEGQSPQITTFQKMPAVTF 60
QY 61 ELTGERDNI FVIEREGGLYYNRALDRETSSTNHLQVAAIDANGIYEGVPITIEYKDN 120
Db 61 ELTGERDNI FVIEREGGLYYNRALDRETSSTNHLQVAAIDANGIYEGVPITIEYKDN 120
QY 121 DNRPTLQSKYEGSVQRNSRPGKPLYYNATDLDDEPATNGQLYYQIVIQLEMINNVMTF 180
Db 121 DNRPTLQSKYEGSVQRNSRPGKPLYYNATDLDDEPATNGQLYYQIVIQLEMINNVMTF 180
```

Db 121 DNRPTFLOSKYEGSVQRNSRPGKPELVYNATLDDPATPNQLYYOIVIQLEPMINMYF 180
Qy 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Db 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Qy 241 APRKVEVENSSTDHPRIKIQVRNMDGQAQSYLVDEKELPRFPFSIDQEGDIYVTOPLDR 300
Db 241 APRKVEVENSSTDHPRIKIQVRNMDGQAQSYLVDEKELPRFPFSIDQEGDIYVTOPLDR 300
Qy 301 EKKDAYFVAVAADDEYKPSYPLEIHVKVDINDNPTCPSPVTFVEVQENRNLGNSIG 360
Db 301 EKKDAYFVAVAADDEYKPSYPLEIHVKVDINDNPTCPSPVTFVEVQENRNLGNSIG 360
Qy 361 TLTHADDEENTANSFLNRYIVEQTPKLPMDGLIQYAGMLQAKOSLKKODTPQYNL 420
Db 361 TLTHADDEENTANSFLNRYIVEQTPKLPMDGLIQYAGMLQAKOSLKKODTPQYNL 420
Qy 421 TIEVSDKDFKLCFVQINVIDINDQIPFEKSDYGNLTAEADNIGSTIITLQATDADEP 480
Db 421 TIEVSDKDFKLCFVQINVIDINDQIPFEKSDYGNLTAEADNIGSTIITLQATDADEP 480
Qy 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Db 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Qy 541 FGKYNASSPAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTVGNVTADEPGLDISY 600
Db 541 FGKYNASSPAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTVGNVTADEPGLDISY 600
Qy 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Db 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Qy 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Db 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Qy 721 INGHARLSTRHTDPEERAAVVLIRINDGRPPLGIVSLVPTFCSCVEGSCFPAGHOT 780
Db 721 INGHARLSTRHTDPEERAAVVLIRINDGRPPLGIVSLVPTFCSCVEGSCFPAGHOT 780
Qy 781 GIPVGAAGVILITLIVIGIILAVFIRIKKQKGNVESASQASEVKPLRS 832
Db 781 GIPVGAAGVILITLIVIGIILAVFIRIKKQKGNVESASQASEVKPLRS 832

RESULT 2
US-08-463-345-1
Sequence 1, Application US/08463345
Patent No. 5710018
GENERAL INFORMATION:
APPLICANT: Anne H. Dantzig, et al.
TITLE OF INVENTION: Mammalian Influx Peptide
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,345
FILING DATE: 05-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/013,462
FILING DATE: 04-FEB-1993
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 832 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-345-1

Query Match 100.0%; Score 4321; DB 1; Length 832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MILQAHLSCLMLATGYGQEGKFSGLKMTFSIYEGQPSQIIFQFKANPAVTF 60
Db 1 MILQAHLSCLMLATGYGQEGKFSGLKMTFSIYEGQPSQIIFQFKANPAVTF 60
Qy 61 ELTGEDNIFVIREGLYYNRALDBETSTHNLQVAALDANGIIVEGVPITIEVDIN 120
Db 61 ELTGEDNIFVIREGLYYNRALDBETSTHNLQVAALDANGIIVEGVPITIEVDIN 120
Qy 121 DNRPTFLOSKYEGSVQRNSRPGKPELVYNATLDDPATPNQLYYOIVIQLEPMINMYF 180
Db 121 DNRPTFLOSKYEGSVQRNSRPGKPELVYNATLDDPATPNQLYYOIVIQLEPMINMYF 180
Qy 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Db 181 QINNKTAISLTRGSOELNPAKPSYNLVISYKMGQSGNSFSDTSDVIITENIMK 240
Qy 241 APRKVEVENSSTDHPRIKIQVRNMDGQAQSYLVDEKELPRFPFSIDQEGDIYVTOPLDR 300
Db 241 APRKVEVENSSTDHPRIKIQVRNMDGQAQSYLVDEKELPRFPFSIDQEGDIYVTOPLDR 300
Qy 301 EKKDAYFVAVAADDEYKPSYPLEIHVKVDINDNPTCPSPVTFVEVQENRNLGNSIG 360
Db 301 EKKDAYFVAVAADDEYKPSYPLEIHVKVDINDNPTCPSPVTFVEVQENRNLGNSIG 360
Qy 361 TLTHADDEENTANSFLNRYIVEQTPKLPMDGLIQYAGMLQAKOSLKKODTPQYNL 420
Db 361 TLTHADDEENTANSFLNRYIVEQTPKLPMDGLIQYAGMLQAKOSLKKODTPQYNL 420
Qy 421 TIEVSDKDFKLCFVQINVIDINDQIPFEKSDYGNLTAEADNIGSTIITLQATDADEP 480
Db 421 TIEVSDKDFKLCFVQINVIDINDQIPFEKSDYGNLTAEADNIGSTIITLQATDADEP 480
Qy 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Db 481 FTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVIIKKPLDEFTAAVSNIVFKANPEPLV 540
Qy 541 FGKYNASSPAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTVGNVTADEPGLDISY 600
Db 541 FGKYNASSPAKFTLIYTDVNEAPOSQHVFOAKVSEDAVIGTVGNVTADEPGLDISY 600
Qy 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Db 601 SLRGDTGMLKIDHVTGEIFSVAPLDRAGSPYRVQVATEVGGSSLSVSEFHLILMDV 660
Qy 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Db 661 NDNPRLAKDYTGIFCHPLSAPGSLIFEATDDQHLFRGPHFTFSIGSSGLQNDMEVSK 720
Qy 721 INGHARLSTRHTDPEERAAVVLIRINDGRPPLGIVSLVPTFCSCVEGSCFPAGHOT 780
Db 721 INGHARLSTRHTDPEERAAVVLIRINDGRPPLGIVSLVPTFCSCVEGSCFPAGHOT 780
Qy 781 GIPVGAAGVILITLIVIGIILAVFIRIKKQKGNVESASQASEVKPLRS 832
Db 781 GIPVGAAGVILITLIVIGIILAVFIRIKKQKGNVESASQASEVKPLRS 832

RESULT 3
US-08-474-067-6

Sequence 6, Application US/08474067
Patent No. 581518
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,067
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-067-6

Query Match 16.1%; Score 696; DB 2; Length 913;
Best Local Similarity 26.9%; Pred. No. 6,9e-53;
Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

QY 126 FLOSKYGVSRQNSRPEKPFYLYNATLDDPATPNGLYQIYQLPMINNMATFOIN--183
DB 41 FPDVHSAVVSRSVHGQPLNVRFGSCDE-----NRKIYFGSSEP 81
QY 184 -----NKTGALISLTREGQELNPAKNSYLVISVKDMGQS-----ENSPEDT 228
DB 82 EDRFVGEDGVVYAKRSQLSAEPE-----FVVSARKETOEKQMKVTKLTPPATFGAS 136
QY 229 SVDIIVTENT-----WKAPKPVEMVNSTDPHPIKITOVWRMDPGA 269
DB 137 EKQOKKIEDIIFPMQYKQSSHLKQKQDWIIP-PINLPENSGPFFQELVRR-----189
QY 270 QYSLVNDKEKLPREFSFI-----DQ-----EGDIYVTOPLDREEDAVYFAVAK 313
DB 190 -----SDDKSLSRYSVTGPGADQPTGIFINPISQSLSVTRKPLREQQLASPLRAHAV 245
QY 314 DEYKPLSYPLKTHVKVDINDNPPCPSPVTVFEVQENRSLGNSIGTLTAHREDENTA 373
DB 246 DVANGQVENIDIVINIDANDNRPELHVYMGTVESGKPGIYVWTVTAIADDPNAQ 305
QY 374 NSFLNTRIYEQTPKLPMDGLFLIQTAYAGMLQALAKQSLKQDTPQYMLTIEVSDKD-----428
DB 306 NGMLRYRILSQAPSSPSPMFTINNETGDIITVAAGLDREKVOQYTLIIQATDMEGNPTY 365
QY 429 -FKTLCEVOINVINDINDQIPF-EKSDYKMLTAEEDINISITLITQATADDPFGSSK 486
DB 366 GLSNATATVITVDVNDNPEFTAMTFYGEV---PENRVDVIVANLITVDKQDPHTPAMN 422

QY 487 ILVHIKGSSEGLGVDTPHNTGVYIIKKPLDFETAAVSNIVFAENBEPVFGVKYN 546
DB 423 ARYQMGGDPTGQFTLLTDBNSNDGLVTVVKPIDFETNRFVLTVAENQVLAKGIQHP 482
QY 547 ASSFAKTLIVTVNNAPOGSHVFOAKVSEDAVIGTKGNVTAQD---EGLDISYSL 602
DB 483 POSTATVSIIVIVNNSPYFVPNPKLVROEGGLASMLTTFARBDRYMOQTSLRYSK 542
QY 603 RGTGRWLKIDHTGELFSVAPLDREA---GSPYVQVATVGGSSLSVSEFHLIM 658
DB 543 LSDPAMMLKIDPVNGQITTTAIVLDRESIYQNNMYNATFLASDNGIIPMSGTGTLOLYL 602
QY 659 DVANDNPPRL-AKDYTGFFCHPLSAPGSLFEATDDQHLFRGPHFTFSISG--SLOND 715
DB 603 DINDNAPQVNPKRAT---TCETIQ-ENAINITAVDDIDPNAGP-PAFELPDSPPSTIKRN 657
QY 716 WEVSKINGTHARLSTHTDEERAYVVLIRINDGAPPLEGIVSLPVTFCSC-VEGSCFR 774
DB 658 WTVIRISGHAOLSLRIFLEAGIYDVPYITDSGNPHASSTSVLKVKVCQCDINGDC-- 715
QY 775 PACHQIGIPV---GNVAGILLTLLVIGITLAVVPIRIKKQKGNVESAQASEVKP 829
DB 716 -----TDVDRIVGAGLGTGATIAIILCIITLILVLMFVMMKRRDKERQAKQLIDP 768

RESULT 4
US-08-474-068A-6
Sequence 6, Application US/08474068A
Patent No. 5837525
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,068A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1683
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-068A-6

Query Match 16.1%; Score 696; DB 2; Length 913;
Best Local Similarity 26.9%; Pred. No. 6,9e-53;
Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

```

QY 126 FLOSKYEGSVRONSREPKFELYVNAATLDDPATPNQOLYYQIVQLPMINWVYFOIN-- 183
DB 41 FPEVDHSAVVSRSVHGQGPLNVRFGSCDE-----NRKIYFGSSSEP 81
QY 184 -----NKTGAISLTREGSQELNPAKNPSYNLVISVKMGQOS-----ENSFSDT 228
DB 82 EDRVGEDGVVAERSFQLSAEPT-----FVVSARKETQEBQMKVKLTPEPAFTGAS 136
QY 229 SVDIIVTENT-----WKAPKPEVEMENSTDPHPRIKITQVRMNDPGA 269
DB 137 EKQOKKIEDIIFPMQYKXSSHLKROKRDWVIR-PINLENSGRPFQOELVRIR----- 189
QY 270 QYSLVDEKXLPREFPSI-----DQ-----EGDIYVTOPLDREKDAVYFAVAK 313
DB 190 -----SDRDKSLSLRYSVTGPGADQPTGIFINPISQSLSTKPLDREKQIASFHLRAHAV 245
QY 314 DEYKPLSYPLEIHVKKDINDNPTCPSPVTVFEVQENBERLGNISGTLTAHREBENTA 373
DB 246 DVNGNOVENPIDIVINIDMNDNRPFLHQVWNGTVEGSKPQTVVMTVAIDADDPNAQ 305
QY 374 NSFLNRYIVEQTPKLPMDGLFLIQTYAGMLQAKOSLKKODTPQVNLITTEVSKD----- 428
DB 306 NGMLRYRILSQAPSSPSPMNETINNEDGITTVAGLDRKVOQYTLIIQATDMEGNPTY 365
QY 429 -FKTLCFVQINVIDINDQIPF-EKSDYGNLTLAEDTNIGSTLITQATDADEPFTGSSK 486
DB 366 GLSNTATAVITVDVNDNPEFTAMTFYGEV---PENRVDIVANLTVTDKQDPTPAMN 422
QY 487 ILVHIHIGDSEGRGLGVDTPHTNTGYIILKKPLDFTFAVSNIVFKAENPEPLVFGVKN 546
DB 423 ARYOMTGDDPTGQFTIITDPSNGLTVVAKPIDFETNRMPVLVALENVPLAKGIQHP 482
QY 547 ASSFAKFTLIVTVNEAPQOSQHVFOAKVSEDAVIGTKVGNVAKDP---EGLDISYSL 602
DB 483 POSTATVSIIVIDVNESPFYVNPVKLVROEBGLAGSMLTFTARDDRVYQOTSLSYK 542
QY 603 RGDTRGMLKIDHVTGEIFSVAPLDREA---GSPRYQVAVTEVSGSSLSVSFHLILM 658
DB 543 LSDPANMLKIDPVNGQITTTAVLDRESIYVQNNMNAATFLASDNGIIPMSGTGLQIYLL 602
QY 659 DVNDNPPRL-AKDYTGFLFCHPLSAPGSLFEATDDQHLFRGPHFTFSLSG--SLQND 715
DB 603 DINNNAQVNPKEAT---TCETLQ-PNAINITAVDPDIDPNAQ-PAFELPDSPPSIKRN 657
QY 716 WEYSKINGTARLSTRHTDEERAYVVLIRINDGRRPLBGIVSLPVTFGSC-VEGSCFR 774
DB 658 WTVIRISGDHQLSLRIRFLEAGIYDVPIVITDSGNPLASSTVLYKVCOCODINDGC-- 715
QY 775 PACHQOTQIPV---GMAVGLITLVLVIGILLAVVIRIKKDKKDNVESQASEVRP 829
DB 716 ----TVDVIRVAGLGTGAILIILCIITLILVLMFVMMKRDYEROKQILLDP 768

```

RESULT 5
US-08-472-481-5
Sequence 5, Application US/08472481
Patent No. 5863804
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,481
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1686
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 913 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-472-481-5

```

Query Match 16.1%; Score 696; DB 2; Length 913;

Best Local Similarity 26.9%; Pred. No. 6,9e-53; Matches 209; Conservative 117; Mismatches 328; Indels 124; Gaps 22;

```

QY 126 FLOSKYEGSVRONSREPKFELYVNAATLDDPATPNQOLYYQIVQLPMINWVYFOIN-- 183
DB 41 FPEVDHSAVVSRSVHGQGPLNVRFGSCDE-----NRKIYFGSSSEP 81
QY 184 -----NKTGAISLTREGSQELNPAKNPSYNLVISVKMGQOS-----ENSFSDT 228
DB 82 EDRVGEDGVVAERSFQLSAEPT-----FVVSARKETQEBQMKVKLTPEPAFTGAS 136
QY 229 SVDIIVTENT-----WKAPKPEVEMENSTDPHPRIKITQVRMNDPGA 269
DB 137 EKQOKKIEDIIFPMQYKXSSHLKROKRDWVIR-PINLENSGRPFQOELVRIR----- 189
QY 270 QYSLVDEKXLPREFPSI-----DQ-----EGDIYVTOPLDREKDAVYFAVAK 313
DB 190 -----SDRDKSLSLRYSVTGPGADQPTGIFINPISQSLSTKPLDREKQIASFHLRAHAV 245
QY 314 DEYKPLSYPLEIHVKKDINDNPTCPSPVTVFEVQENBERLGNISGTLTAHREBENTA 373
DB 246 DVNGNOVENPIDIVINIDMNDNRPFLHQVWNGTVEGSKPQTVVMTVAIDADDPNAQ 305
QY 374 NSFLNRYIVEQTPKLPMDGLFLIQTYAGMLQAKOSLKKODTPQVNLITTEVSKD----- 428
DB 306 NGMLRYRILSQAPSSPSPMNETINNEDGITTVAGLDRKVOQYTLIIQATDMEGNPTY 365
QY 429 -FKTLCFVQINVIDINDQIPF-EKSDYGNLTLAEDTNIGSTLITQATDADEPFTGSSK 486
DB 366 GLSNTATAVITVDVNDNPEFTAMTFYGEV---PENRVDIVANLTVTDKQDPTPAMN 422
QY 487 ILVHIHIGDSEGRGLGVDTPHTNTGYIILKKPLDFTFAVSNIVFKAENPEPLVFGVKN 546
DB 423 ARYOMTGDDPTGQFTIITDPSNGLTVVAKPIDFETNRMPVLVALENVPLAKGIQHP 482
QY 547 ASSFAKFTLIVTVNEAPQOSQHVFOAKVSEDAVIGTKVGNVAKDP---EGLDISYSL 602
DB 483 POSTATVSIIVIDVNESPFYVNPVKLVROEBGLAGSMLTFTARDDRVYQOTSLSYK 542
QY 603 RGDTRGMLKIDHVTGEIFSVAPLDREA---GSPRYQVAVTEVSGSSLSVSFHLILM 658
DB 543 LSDPANMLKIDPVNGQITTTAVLDRESIYVQNNMNAATFLASDNGIIPMSGTGLQIYLL 602
QY 659 DVNDNPPRL-AKDYTGFLFCHPLSAPGSLFEATDDQHLFRGPHFTFSLSG--SLQND 715
DB 603 DINNNAQVNPKEAT---TCETLQ-PNAINITAVDPDIDPNAQ-PAFELPDSPPSIKRN 657

```


TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-643-42

Query Match 15.8%; Score 684.5; DB 1; Length 916;
Best Local Similarity 29.3%; Pred. No. 7.3e-52;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKEVENENSTDPPIKITOVW---NDPAQYSL--VDKEKLPRPPSID-OEGDI 292
171 WVIP-PINPENSNGRFPQOLVIRSDKNDIPRISITGVADQPEMEVFSINSMGRM 229
293 YTPQPLDREKDAVYFAVAKDEYKPLSYPLHVKVDINDNPPCSPVTFVEQEN 352
230 YVTRPMDREHASYHRAHADVMDNKNKVENPIDLYIVIDMNDHPEFIHQVNCSDVEG 289
353 ERGNSIGTLTAHREENTANSFLNRYIVEQPKLPMGLFLIQYAGMLQAKOSLKK 412
290 SKRGTYVMTITANDADSTANGVNRIRYVOTPOSSQOMFTINSRTGDIIVVAAGWDR 349
413 QDPQYVNLTEVSDKD-----FKTLCFVOINVINDIQPIPEKSDYGNLTAEPTNIG 466
350 EKVOQYVIVVQATDMGNLNYGLSNTATLITVDVNDNSEFTASTFAG--EVPENSVE 407
467 STLTITQATDADEPFTGSSKILYHIKGDSEGRGVDTDPHTNNGYVYIKKPLDFTAAV 526
408 TVVANLTVMDRDPHSPNNNAVYRIISGDPGHSFVRTDPVTNEGAVTVKADYELNRA 467
527 SNIVFAKENPEPLVFGKYNASSPAKFTLLVTVDNENAPPOSHVQAKVSEDAVAIGTKVG 586
468 FMLTVMVSNQAPLASGIQMSFGSTAGTISIMDINEAPYPSNKKLRLEEGVPGTVLT 527
587 NVTAKDEGL---DISYSLRGDTRGWLKIDHVTGEIFSVAPLDREA---GSPYRQOVA 639
528 TFSADVDPDRMQAVRYSKLSDPASMLHINATNGQITTVAVLDRESLYTKNNYEATFLA 587
640 TEVGGSSLSVSEFHLIMDVNDNPPRLADYGLFCHPLSAGSLIFEATDDOHLFR 699
588 ADNGIPPASGTGLQIYLIDINDNAPLPEKAOICERPMLNA---INTTAADADVHPNI 644
700 GPH-FPTSLSGSLQNDMEVSKINGTHARLSTHTDEERAYVVLIRINDGRPLEGIV 758
645 GPYVFEPLFPVPAVRKMTITRLNGDVAQLSLILYLEAGMYDVPIITVDSGNPLSNTS 704
759 SLPTFSCVYEGSCFRPAGHOTGIPYGMVAGILLTLLVIGIILAV--FIRIKDKGX 816
705 IIKKVCPCDNDGCTTIG---AVAAAGLGTGAIVAILICILILLTWLLFVMMKKEREX 761
QY 817 D 817
DB 762 E 762

RESULT 8
US-08-332-638-48
Sequence 48, Application US/08332638
Patent No. 5646250
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,638
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,643
FILING DATE: 17 APR 1992
APPLICATION NUMBER: US/08/049,460
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5646250and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31340
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 916 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-638-48

Query Match 15.8%; Score 684.5; DB 1; Length 916;
Best Local Similarity 29.3%; Pred. No. 7.3e-52;
Matches 176; Conservative 106; Mismatches 288; Indels 31; Gaps 12;

239 WKAPKEVENENSTDPPIKITOVW---NDPAQYSL--VDKEKLPRPPSID-OEGDI 292
171 WVIP-PINPENSNGRFPQOLVIRSDKNDIPRISITGVADQPEMEVFSINSMGRM 229
293 YTPQPLDREKDAVYFAVAKDEYKPLSYPLHVKVDINDNPPCSPVTFVEQEN 352
230 YVTRPMDREHASYHRAHADVMDNKNKVENPIDLYIVIDMNDHPEFIHQVNCSDVEG 289
353 ERGNSIGTLTAHREENTANSFLNRYIVEQPKLPMGLFLIQYAGMLQAKOSLKK 412
290 SKRGTYVMTITANDADSTANGVNRIRYVOTPOSSQOMFTINSRTGDIIVVAAGWDR 349
413 QDPQYVNLTEVSDKD-----FKTLCFVOINVINDIQPIPEKSDYGNLTAEPTNIG 466
350 EKVOQYVIVVQATDMGNLNYGLSNTATLITVDVNDNSEFTASTFAG--EVPENSVE 407
467 STLTITQATDADEPFTGSSKILYHIKGDSEGRGVDTDPHTNNGYVYIKKPLDFTAAV 526
408 TVVANLTVMDRDPHSPNNNAVYRIISGDPGHSFVRTDPVTNEGAVTVKADYELNRA 467
527 SNIVFAKENPEPLVFGKYNASSPAKFTLLVTVDNENAPPOSHVQAKVSEDAVAIGTKVG 586
468 FMLTVMVSNQAPLASGIQMSFGSTAGTISIMDINEAPYPSNKKLRLEEGVPGTVLT 527
587 NVTAKDEGL---DISYSLRGDTRGWLKIDHVTGEIFSVAPLDREA---GSPYRQOVA 639
528 TFSADVDPDRMQAVRYSKLSDPASMLHINATNGQITTVAVLDRESLYTKNNYEATFLA 587
640 TEVGGSSLSVSEFHLIMDVNDNPPRLADYGLFCHPLSAGSLIFEATDDOHLFR 699
588 ADNGIPPASGTGLQIYLIDINDNAPLPEKAOICERPMLNA---INTTAADADVHPNI 644
700 GPH-FPTSLSGSLQNDMEVSKINGTHARLSTHTDEERAYVVLIRINDGRPLEGIV 758
645 GPYVFEPLFPVPAVRKMTITRLNGDVAQLSLILYLEAGMYDVPIITVDSGNPLSNTS 704
759 SLPTFSCVYEGSCFRPAGHOTGIPYGMVAGILLTLLVIGIILAV--FIRIKDKGX 816
705 IIKKVCPCDNDGCTTIG---AVAAAGLGTGAIVAILICILILLTWLLFVMMKKEREX 761
QY 817 D 817
DB 762 E 762

RESULT 9
US-08-188-228-62
Sequence 62, Application US/08188228
Patent No. 5597725
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Borun
STREET: 6300 Sears Tower, 233 S. Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,228
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/049,460
FILING DATE: 19 APR 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/872,643
FILING DATE: 17 APR 1992
ATTORNEY/AGENT INFORMATION:
NAME: No. 559725and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31340
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-188-228-62

Query Match 15.2%, Score 657.5; DB 1; Length 713;
Best Local Similarity 28.7%; Pred. No. 1.2e-49;
Matches 199; Conservative 104; Mismatches 285; Indels 105; Gaps 23;

QY 160 NGQLYQIVQLEPMINNVTFQINNKTKGALSLSL--TREGSGELNPAKNPSYN-----LVI 211
DB 61 NDKRLRYEV-----SSPYKXVNSDGLVALRNTAVGKTLFVHARTPHAEDEWAEIV 112
QY 212 SVKMGQSEN--SFSOTTSV-----DIYVENIMKAPKPEVEMENSTDEHPKITTOYRW 264
DB 113 GSKDIQSLQDIFFEARTSPVPROKRSIVS-----PLILENORQPPRPVGCY-- 162
QY 265 NDPAQASLVDEKELPRPF-----SIDOE-----GDIYVQPLDREKDAYVF 308
DB 163 -----VSDPPEKSKRLTKGVQDEPKGIFRINENTGSVSTRIDREVIANYQL 213
QY 309 YAAKDEYKPLSYLEIHHVKVDINDNPTCPSPVTFVEVQENRSLNSIGTLTJARD 368
DB 214 FVEETDVNGKTLIEGPVLEVIDQNNRPFRREGPYIGHWEGSPGTVMRTAFDAD 273
QY 369 EENANSEFLWRYIEOTPKLPMDELFLIQTYAG--MLQAKQSLKQOT---PQYNLTIEV 424
DB 274 DPATDNLALRNKIQOTPKRSPNMFYIDPEKGDIVTVVSPALLDRETLLENPKELIEA 333

QY 425 SDR-----DEKLCFQVQINVIDINDQIPFEKSDYGNLTAEPTNIGSTLTITQATDADE 479
DB 334 QDMAGLDVGLTGATATATIMIDDKNDSPKFTKKEF--QATVEGAVG-VIVNLTVEDXKD 390
QY 480 PFTGSSKILYHIKGBSEGLGVDTPHNTGVIYIKKPLDFTAAVSNVFAENEP 539
DB 391 PFTGAMRAAYTIINGNGQSFELHTNPQTEGMLSVKPLDIYERSAHTLLIKVENDDL 450
QY 540 VFGVKYNASEFAKFTLVTVTNEAPQPSQVFAKVEDVAIGTKVGNATKXDPEDGD-- 597
DB 451 VPDVSYSPSSATYHTVLDVNEGPFYPPPMVATROEDLSVGSVLTVAATPDPSLOH 510
QY 598 -ISVSLRGDTRGMKLDIHTVGEIJSVAPLDREA---GSPRYQVAVTEVGGSSLSVSE 652
DB 511 TIRSYVKDPAQMINIPINGVDTTAVLDRESPVDNSVYTLFLAIDSGNPATGTGT 570
QY 653 PHLLMDVNDNPPRLAKDYGLFPCHPLSAP-----GSLIFEATDDQHLFRGPHFT 704
DB 571 LLITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVILGASDKDLHPNTPD-FK 619
QY 705 FSLGSGSLQND-WEVSKINGTHARLSTRHTDFEBRAYVVLIRINDGRPLEGIVSLPYT 763
DB 620 FEIHKQAVPKVWKISKINTTHALVSLQ--NLNKNANPLIMVTDGSKPPMTNITDLRVQ 678
QY 764 FCSGVBG--SCFRPAGHQTGIPVGMVAVGILLT 794
DB 679 VCSGRNSKVDCAAGALRFLPSV-----ILL 706

RESULT 10
US-08-332-643-56
Sequence 56, Application US/08332643
Patent No. 5639634
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
ADDRESS: Bicknell
STREET: Two First National Plaza, 20 South Clark
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,643
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/872,643
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: No. 5639634and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/30795
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-9740
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 713 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-332-643-56

Query Match 15.2%; Score 657.5; DB 1; Length 713;
 Best Local Similarity 28.7%; Pred. No. 1.2e-49;
 Matches 199; Conservative 104; Mismatches 285; Indels 105; Gaps 23;

160 NGOLYQVIVQLPMINNMVFOINNKGTALSL---TREGSOELNPAKPSYN-----LVI 211
 61 NDKLRREV-----SSPFKNSDGLVALNRTNNGKTLFVHARTPHAEDEMAELVIV 112
 212 SVKMGGOSEN--SFSDTTSV-----DIIVTENIMKAPKEVENENSTDPHPIKITQVRW 264
 113 GSKDIQSLQDIFKFAFRTSPVROKRSIVS-----PILIPENORQFPFDDVGRV-- 162
 265 NDPAQVSLVDKELRPFPP-----SIDOE-----GDIYVQPLDREKDAYVF 308
 163 -----VDSRPERSKRFLTGKGVDPKGIFRINENTGSVSTRITLDREVIAYQL 213
 309 YAAKDEYKPLSLPELHVAKVDINDNPTCPSPVTFVEQENRIGNSIGTLTAHARD 368
 214 FVETTVNGKTLGEPVLEVIYIDQNDRIIFREGPIYIGHMGSPTGTTWMTAFDAD 273
 369 EENTANSFLNRYIVEQPKLPMDGLFLIQTYAG-MLQAKOSLKKQDT---POYNLTIEV 424
 274 DPATDNLALRYNIRQOTPKSPMMFYIDPEKGDIVTVVSPALLDRETLLENPKYELIEA 333
 425 SDK-----DKITLCFQVQINVIDINDQPIFEKSDYGNLTLAEDTNGSTLITQATDAE 479
 334 QDMAGLDVGLTGATATIMIDDKNDSPKTKKEF--QATVEGAVG-VIVNLTVEDKDD 390
 480 PFTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVYIKKPLDFTAAVSNIVFAENDEPL 539
 391 PTTGAMRAAYTIINGNPGOSFEIHTNPOTNEGMLSVKPLDYELISAFHTLLIKVENEDPL 450
 540 VFGVKNASSPAKFTLLIYTDVNEAPQFSQHVFOAKVSEDAVIGTKVGNVTAKDPEGID-- 597
 451 VPDVSYGSSSTAYVHTIYLDVNEGVPYDPDMVTRQEDLSVGSVLLTVNATDSDSIQH 510
 598 -IYSLSRGDTGMLKIDHVTGEIRSVAPLDREA-----GSPRYQVVAEVEGSSLSVSE 652
 511 TIRSYVKDPAGMININPINGTVDTTAVLDRESFVDNSVYTLAFLAIDSGNPPATGTGT 570
 653 FHLIMDVNDNPPRLAKDYGLFCFCHPLSAP-----GSLIFEATDDOHLFRGPHFT 704
 571 LLITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVILGASDKDLHPNTDP-FK 619
 705 FSLGSGSLQND-WEVSKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLEGIVSLPVT 763
 620 FEIKQAVPDVQWYKISKINNTHALVSLIQ-NLNKANYMLPIWTDVSGRPMNTITDRLVQ 678
 764 FCSCVEG--SCFRAGHOTGIPTYGMAVGILLT 794
 679 VCSCRNKVDNCNAGALRFLPSV-----ILLIS 706

RESULT 11

US-08-332-638-62
 Sequence 62, Application US/08332638
 Patent No. 5646250

GENERAL INFORMATION:
 APPLICANT: Suzuki, Shintaro
 TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 STREET: 6300 Sears Tower, 233 S. Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60606
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/332,638
 FILING DATE: 01-NOV-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/872,643
 FILING DATE: 17 APR 1992
 APPLICATION NUMBER: US/08/049,460
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5646250and, Greeta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 31340
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 474-6300
 TELEFAX: (312) 474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 62:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 713 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

US-08-332-638-62

Query Match 15.2%; Score 657.5; DB 1; Length 713;
 Best Local Similarity 28.7%; Pred. No. 1.2e-49;
 Matches 199; Conservative 104; Mismatches 285; Indels 105; Gaps 23;

160 NGOLYQVIVQLPMINNMVFOINNKGTALSL---TREGSOELNPAKPSYN-----LVI 211
 61 NDKLRREV-----SSPFKNSDGLVALNRTNNGKTLFVHARTPHAEDEMAELVIV 112
 212 SVKMGGOSEN--SFSDTTSV-----DIIVTENIMKAPKEVENENSTDPHPIKITQVRW 264
 113 GSKDIQSLQDIFKFAFRTSPVROKRSIVS-----PILIPENORQFPFDDVGRV-- 162
 265 NDPAQVSLVDKELRPFPP-----SIDOE-----GDIYVQPLDREKDAYVF 308
 163 -----VDSRPERSKRFLTGKGVDPKGIFRINENTGSVSTRITLDREVIAYQL 213
 309 YAAKDEYKPLSLPELHVAKVDINDNPTCPSPVTFVEQENRIGNSIGTLTAHARD 368
 214 FVETTVNGKTLGEPVLEVIYIDQNDRIIFREGPIYIGHMGSPTGTTWMTAFDAD 273
 369 EENTANSFLNRYIVEQPKLPMDGLFLIQTYAG-MLQAKOSLKKQDT---POYNLTIEV 424
 274 DPATDNLALRYNIRQOTPKSPMMFYIDPEKGDIVTVVSPALLDRETLLENPKYELIEA 333
 425 SDK-----DKITLCFQVQINVIDINDQPIFEKSDYGNLTLAEDTNGSTLITQATDAE 479
 334 QDMAGLDVGLTGATATIMIDDKNDSPKTKKEF--QATVEGAVG-VIVNLTVEDKDD 390
 480 PFTGSSKILYHIKGDSEGRIGVDTDPHTNTGYVYIKKPLDFTAAVSNIVFAENDEPL 539
 391 PTTGAMRAAYTIINGNPGOSFEIHTNPOTNEGMLSVKPLDYELISAFHTLLIKVENEDPL 450
 540 VFGVKNASSPAKFTLLIYTDVNEAPQFSQHVFOAKVSEDAVIGTKVGNVTAKDPEGID-- 597
 451 VPDVSYGSSSTAYVHTIYLDVNEGVPYDPDMVTRQEDLSVGSVLLTVNATDSDSIQH 510
 598 -IYSLSRGDTGMLKIDHVTGEIRSVAPLDREA-----GSPRYQVVAEVEGSSLSVSE 652
 511 TIRSYVKDPAGMININPINGTVDTTAVLDRESFVDNSVYTLAFLAIDSGNPPATGTGT 570
 653 FHLIMDVNDNPPRLAKDYGLFCFCHPLSAP-----GSLIFEATDDOHLFRGPHFT 704
 571 LLITLEDVNDNAP-----FIYPTVAEVCDDAKNLSVILGASDKDLHPNTDP-FK 619
 705 FSLGSGSLQND-WEVSKINGTHARLSTRHTDPEERAYVVLIRINDGRRPLEGIVSLPVT 763

Db 620 FEIHKQAVPKVKIKSINTHALVSLIQ-NLANKANTNPLMTDSKPEPMTNITDLRVQ 678
QY 764 FCSCVEG--SCFRPAGHOTGIPVGMVAILLT 794
Db 679 VCSGRNSKVDNMAAGALRPLSPV-----ILLS 706

RESULT 12

US-08-453-702A-98
Sequence 98, Application US/08453702A
Patent No. 5891706
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
STREET: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,702A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5891706and, Greta B.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32657
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ. ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 555 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-702A-98

Query Match 15.1%; Score 650.5; DB 2; Length 555;
Best Local Similarity 30.5%; Pred. No. 3,3e-49;
Matches 174; Conservative 87; Mismatches 252; Indels 57; Gaps 13;

QY 239 WKAPKPVENVSTDPHPKIKITOVKRN-----DPGQYSLVDEKELPRPPEST 286
Db 2 WVIP-PINLPENSRGPPELVRIRSRDNLRLSVSVTPGA-----DQPPTGIFIL 53
QY 287 DQ-EGDIYVTPDIDREKDAVYFAVAKDEYKPLSYPLEIHKVKVDINDNPPTCSPTV 345
Db 54 NPISGQLSVTKPDRDLIAFHRAHAVDINGNOVENPIDIVINVDMDNDRPEPTLHQW 113
QY 346 VFEVQENRIGNSIGTLTAHDREENTANPLNRYIVEQTPKLPMDGLFIQTYAGMLQL 405
Db 114 NGSVPEGSRGKGYVMTVTALIDADDPMALNGMLRYRLISQAPSPSPMFMFINNETGDIIT 173
QY 406 AKQSLKKQDTPQVNLITIEVSKD-----FKTLCFQVQINVIDINDQIPF-EKSDYGNLT 458
Db 174 VAAGLREKVOYQTLTIQATDMGKNPTYGSLNATRAVITVTDNDNPREPTAMTFYGEV- 232
QY 459 LAEDTNGISLTITLTQADDEPFTGSSKILYHIKGDSEGRLEGVDPDPHNTGTVIILKRP 518
Db 233 --PENNDVIVVALYTTDQGHPTPAMNAARYISGDDPTGRFALLDPSNDGLVTVVKE 290
QY 519 LDEFJAASNTVKAENPEPLVFGVYKXNSSFATKFTLITVDNEADQFSQHFQAKVSD 578

Db 291 IDEFTNMEFVITVAENQVPLAKGIOHPPOSTATVSVTVADVNEENFYFAPNPKIRQEEG 350
QY 579 VAIGTKVQNTADPEGL---DISYLRGDTROMKLDHYTGRIFFSAPLDRRA-----GS 631
Db 351 LHAQTMULTITADPQRYMOQNIYTKLSDPANWLKIDPNQGITTIIVADRSPYQNN 410
QY 632 PYRQVATEVGGSSLSVSEFHLIMDVNDNPRLAKDYTGJFCHPLSA-----PG 684
Db 411 IVNATPLASNGIIPMSGCTLOIYILDINDNAPVL-----PQAEFCETPEPN 460
QY 685 SLIFEATDDQHLFRGPH-FTFSLGSGSLQNDWEVSKINGTARLSTRHTDPEERAYVL 743
Db 461 SINIAALDYVIDPNAQPFAPDLPLSPVTTIKRMTIRLNGDFAQLMKIKFLAAGIYEVP 520
QY 744 IRINDGRPRLBEGIVSLPTFCSC-VEGSC 772
Db 521 IITDSGNPKPSNISILRVKVCQCDNSGDC 550

RESULT 13

US-08-474-067-7
Sequence 7, Application US/08474067
Patent No. 5811518
GENERAL INFORMATION:
APPLICANT: Ranscht, Barbara
TITLE OF INVENTION: T-Cadherin Adhesion Molecule
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,067
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/213,361
FILING DATE: 14-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/607,293
FILING DATE: 30-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 1682
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 837 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-474-067-7

Query Match 14.7%; Score 634; DB 2; Length 837;
Best Local Similarity 30.4%; Pred. No. 2e-47;
Matches 189; Conservative 104; Mismatches 261; Indels 68; Gaps 25;

QY 239 WKAPKPVENVSTDPHPKIKITOVKRN-----DPGQYSLVDEK-ELKRPPESTIDE-GDI 292
Db 112 WVIP-PISCLNHRGYPRLVQIKSNKDKSKYYSITGQADSPVGIFFIERTGWL 170
QY 293 YVQPLDREKDAVYFAVAKDEYKPLSYPLEIHKVKVDINDNPPTCSPVTFEVQEN 352

THIS PAGE BLANK (uspiu)

BEST AVAILABLE COPY

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 22, 2003, 01:06:35 ; Search time 435 Seconds
(without alignments)
4307.270 Million cell updates/sec

Title: US-10-025-380-1081
Perfect score: 4321
Sequence: 1 MIOQAHSLCLMLYLATG.....DKGKNVESQAQSEVKEFLRS 832

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=Cgmr2_1/USPTO.spool/US10025380/rnatc_17062003_173327_24656/app.query.fasta_1.1031
-DB=N.geneseq_101002 -QFMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PCT -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10025380 @CGN 1.1 290 @rnatc_17062003_173327_24656 -NCPU=6 -ICPU=3
-NO MMAP -IARBOQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOCG
-DEF_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.GeneSeq_101002:*

1:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1980.DAT:*
2:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1981.DAT:*
3:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1982.DAT:*
4:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1983.DAT:*
5:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1984.DAT:*
6:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1985.DAT:*
7:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1986.DAT:*
8:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1987.DAT:*
9:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1988.DAT:*
10:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1989.DAT:*
11:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1990.DAT:*
12:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1991.DAT:*
13:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1992.DAT:*
14:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1993.DAT:*
15:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1994.DAT:*
16:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1995.DAT:*
17:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1996.DAT:*
18:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1997.DAT:*
19:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1998.DAT:*
20:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA1999.DAT:*
21:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2000.DAT:*
22:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2001A.DAT:*
23:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2001B.DAT:*
24:	/SIDS2/gcgdata/geneSeq/geneSeq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4321	100.0	3345	20	AA18166	Human HPT-1 protei
2	4321	100.0	3345	22	AA129511	Human intestinal p
3	4321	100.0	3345	24	ABN6791	Gene #3289 used to
4	4321	100.0	3345	24	AB162444	Colon adenocarcino
5	4321	100.0	3345	24	AB164424	Stomach cancer rel
6	4303	99.6	3654	24	ABN97312	Gene #3810 used to
7	4303	99.6	3654	24	AB162484	Colon adenocarcino
8	4303	99.6	3654	24	AB167667	Oesophagus cancer
9	922.5	21.3	2820	22	AA198190	Human EST-derived
10	922.5	21.3	2820	22	AA157507	Human kidney cell
11	908.5	21.0	2848	21	AA137087	Human PRO1340 (UNQ
12	908.5	21.0	2848	22	AA192106	Human PRO1340 cDNA
13	908.5	21.0	2848	22	AA154364	DNA encoding prote
14	908.5	21.0	2848	24	ABK33629	cDNA encoding huma
15	878.5	20.3	2938	23	AA155602	Human cDNA encodin
16	771.5	17.9	3540	22	AA194573	Human full-length
17	758.5	17.6	5316	24	ABK44858	cDNA encoding colo
18	699.5	16.2	2853	24	ABN89391	Human N-cadherin e
19	699.5	16.2	4136	23	AA178503	DNA encoding novel
20	696.5	16.1	2824	24	ABN89350	Chicken N-cadherin
21	696.5	16.1	2824	24	ABN89351	Chicken N-cadherin
22	694.5	16.1	3875	12	AA011561	Partial endothelia
23	694.5	16.1	4321	24	AB199593	Mouse ischaemic co
24	684.5	15.8	3048	18	AA178543	Human cadherin-4 c
25	684.5	15.8	3048	18	AA161920	Human cadherin-4 c
26	684.5	15.8	3048	24	ABK83876	Human cDNA differe
27	671.5	15.5	3551	15	AA072597	Human HT-1376 cell
28	671.5	15.5	3552	20	AA224623	Human lung tumor a
29	671.5	15.5	3552	21	AA165862	Human lung cancer-
30	671.5	15.5	3552	24	AB149081	Human lung tumour
31	671.5	15.5	3742	22	AA1998127	Human late stage o
32	667.5	15.4	3212	24	AB166329	Lung cancer relate
33	667.5	15.4	4125	24	AA158391	Human polynucleoti
34	665.5	15.4	3364	22	AA160177	Human polynucleoti
35	660.5	15.3	2690	18	AA178545	Human cadherin-13
36	660.5	15.3	2690	18	AA178545	Human cadherin-13
37	630.5	14.6	3279	13	AA024857	Full length human
38	630.5	14.6	3959	13	AA024856	Sequence of T-cadh
39	624.5	14.5	3156	23	AA186134	DNA encoding novel
40	611	14.1	14530	24	AA199765	Rat fat 3 protein
41	605	14.0	13960	23	AA187164	DNA encoding novel
42	604	14.0	2768	24	ABN89352	Mouse E-cadherin e
43	604	14.0	2768	24	ABN89353	Mouse E-cadherin e
44	601.5	13.9	12791	23	AB111557	Drosophila melanog
45	599.5	13.9	17282	23	AB111556	Drosophila melanog

ALIGNMENTS

RESULT 1
ID AAX18166 standard, DNA, 3345 BP.
XX AAX18166;
AC
XX
DT 04-MAY-1999 (first entry)
XX
DE Human HPT-1 protein coding sequence.
XX
KW Gastro-intestinal transport receptor; binding protein; hsi; HPT1;
KW DH; hPEPT1; human; GI tract receptor; sucrose-isomaltase complex;
KW intestinal peptide-associated transporter; hypertension; diabetes;
KW osteoporosis; haemophilia; anaemia; cancer; migraine; angina pectoris;
KW therapeutic agent delivery; therapy; probe; ss.
XX
OS Homo sapiens.
XX

PN M09851325-A2.
 XX 19-NOV-1998.
 XX 15-MAY-1998; 98WO-US10088.
 XX 15-MAY-1997; 97US-0046595.
 PR (CYTO-) CYTOGEN CORP.
 PA (ELAN-) ELAN CORP PLC.
 XX Alvarez VL, Belinka BA, Cagney GM, Carter JM, Lambkin LV;
 PI Omahony DJ, Patterson CA, Singleton J;
 DR WPI; 1999-009566/01.
 DR P-PSDB; AAW74089.
 XX
 PT New proteins that bind specifically to receptors in the
 PT gastro-intestinal tract and related nucleic acid - chimaeras and
 PT antibodies, used to deliver therapeutic or diagnostic agents to, or
 PT through, the gastrointestinal tract, e.g. insulin or leuprolide
 XX
 PS Disclosure; Fig 2; 294pp; English.
 XX
 CC This sequence encodes the human HPT-1 protein. The invention relates to
 CC purified proteins (I) that bind specifically to at least one of the
 CC gastro-intestinal (GI) tract receptors human intestinal
 CC peptide-associated transporter (HPT1), hPPT1, D2H and human
 CC sucrase-isomaltase complex (hSIC). (I) provide active transport of
 CC therapeutic agents through human and animal GI tissue (into the blood)
 CC for in vivo delivery, particularly for treatment or prevention
 CC of hypertension, diabetes, osteoporosis, haemophilia, anaemia, cancer,
 CC migraine, or angina pectoris. Specifically they are used to deliver
 CC insulin or leuprolide, but many other suitable therapeutic agents are
 CC disclosed, including genes or inhibitory nucleic acid, imaging agents and
 CC antigens. (II) may also provide targeting to the GI tract. Other uses of
 CC (I) are: (i) to determine the level of specified receptors in a sample
 CC (in a binding assay); and (ii) to screen for molecules that bind (I).
 CC Immunogenic analogues or derivatives of (I) are used to raise antibodies
 CC and in immunoassays. The antibodies are used to locate, detect and
 CC measure (I), e.g. for imaging, monitoring treatment, tissue analysis
 CC etc., also for peptide purification and immobilisation.
 CC
 XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-10-025-380-1081 (1-832) x AAX18166 (1-3345)
 QY 1 MetIleuGlnAlaHisIleuHisSerIeuCysIeuLeuMetIeuTyrLeuAlaThrGly 20
 DB 88 ATGAACTTCAGGCCCATTCACCTCCGTGCTCTTATGCTTATTTGGCACTGGA 147
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysPrometThrPheSerIleTyrGlu 40
 DB 148 TATGGCCCAAGGGGGAAGTTTAACTGACCCCTGAACCCATGACATTTCTATTATTA 207
 QY 41 GlyGlnIupProSerGlnIleIlePheGlnPheLysAlaAsnProAlaValIThrPhe 60
 DB 208 GGCCCAAGAACCGAGTCAATATATTCAGATTAAAGCCCAATCCCTCGCTGTGACTTT 267
 QY 61 GluLeuThrGlyIuThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80
 DB 268 GAACCTAACTGGAGACACACACATATTTGTGATGAAGACGGAGGACCTTGTATTAC 327
 QY 81 AsnArgAlaLeuAspArgIuThrArgSerThrHisAsnLeuGlnValAlaLeuAsp 100

DB 328 AACAGAGCTTGGACAGGGAAACAGATCTACTACATCTCCAGGTGGACCCCTGGAC 387
 QY 101 AlaAsnGlyTyrIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB 388 GCTAATGGAATTTATGAGAGGGTCCAGTCCCTATACCACTAAGATGAAGCATCTAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGluGlySerValArgGlnAsnSerArg 140
 DB 448 GACATGACCCCAAGTTCTCCAGTCAAGTACGAGGGCTCAGTAAAGGCAACTCTCGC 507
 QY 141 ProGlyLysProPheLeuTyrValAlaAsnIleThrAspLeuAspAspProAlaThrProAsn 160
 DB 508 CCAGAAAGCCCTTCTGTGTATGTCATGACACAGACCTGGATGATCCGGCACTCCCAAT 567
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 DB 568 GGCCAGCTTTATTAACAGATGTCATCCAGCTTCCATGATCAACATGTCATGACTTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnIuLeuAsn 200
 DB 628 CAGATCAACAACAAACGGGAGCCATCTCTTACCAGAGGAGATCTCAGAAATGAAT 687
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
 DB 688 CTGCTAAGAAATCCTTCTTATATCTGGTATCTCAGTGAAGACATGGAGGCCAGAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValIThrGluAsnIleTyrLys 240
 DB 748 GAGAAATCTTCAGTGAATACCAATCTGTGATATCTAGTACAGAGAAATATTGGAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr 260
 DB 808 GCACCAAAACCTGTGAGAGATGGGAAACCTCAACAGATCTCCACCCCAACAAATCACT 867
 QY 261 GlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLysGlyLysPro 280
 DB 868 CAGGTGGGTGGATGATCCCGGTGACAAATATCTCTGATTACAAAGAAAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValIThrGlnProLeuAspArg 300
 DB 928 AGATTCCTCATTTCAATGACCAAGAAAGATATTTACGCTCAGCCCTTGGACCGA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 DB 988 GAGAAAGAGATGCATATGTTTTTTATGCACTTGAAGATGATGACGAAACCACTT 1047
 QY 321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 DB 1048 TCATATCCGCTGGAAATTCATGTAAAGATTATATATATATATCCACCTACATGT 1107
 QY 341 ProSerProValIThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB 1108 CCGTCACACGTAAACCGATTGAGGTCCAGAGAAATGAACGATGGGTAAACGATTGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluLysAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 DB 1168 ACCCTTATGCAATGACAGAGATGAAGAAATATCCGCAACGTTTCTTAATCAAGC 1227
 QY 381 IleValGluGlnThrProLysLeuProMetArgGlyLeuPheLeuIleGlnThrTyrAla 400
 DB 1228 ATTGTGAGCAAACTCCCAAACTTCCATGATGACATCTCTTCAAAACCAACTAAGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysLysGlnAspThrProGlnTyrAsnLeu 420
 DB 1288 GGAATGTAAAGTTAGTAAACAGTCTTGAAGAGCAAAATCTCTCGTACAACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 DB 1348 ACAGTAAAGGTGTCTGCAAAAGATTCACAAACCTTTGTTGTGCAAAATCAAGTTATT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAla 460
 DB 1408 GATATCAATGATCAATGCCATCTTGAAGAAATTCAGATTATGAAACCTGACCTTGTCT 1467

QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB 1468 GAAGACACAAACATGGGGTCCACCATCTTAACCATCCAGGCCACTGATGCTGATGAGCCA 1527
 QY 481 PheThrGlySerSerIleValLeuThrHisIleIleIleValSerGlyAspSerGluValLeu 500
 DB 1528 TTTTACTGGGAGTTCTTAAATTCGTATCATATCAATAAGGAGACACTGAGGAGCGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleIleValSerProLeuAsp 520
 DB 1588 GGGGTTACACAGATCCCATCCACACACCGGATATCTCATATTAATAAGCCTCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuVal 540
 DB 1648 TTTGAACACACAGCTGTTTCCACATGTGTTCAACACAGAAATCTGAGCCTCTAGTG 1707
 QY 541 PheGlyValAlaValPheThrAsnAlaSerSerPheAlaValPheThrLeuIleValThrAspVal 560
 DB 1708 TTTGGTGTGAAGTACAAATGCAAGTCTTTTGGCCAAAGTCCATTAATTTGACAGATGTG 1767
 QY 561 AsnGlyAlaProGluPheSerGlnHisValPheGlnAlaValSerGluAspValAla 580
 DB 1768 AATGAAGCACTCAATTTTCCCAACAGTATTTCCAAAGCGAAAGTCACTGAGATGTAGCT 1827
 QY 581 IleGlyThrIleValGlyAsnValThrAlaValAspProGluGlyLeuAspIleSerTyr 600
 DB 1828 ATAGGCATTAAGTGGGCAATGTGACTGCCAAGATCCAGAAAGTCTGCACTATAGCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLeuValIleAspHisValThrGlyIleIlePhe 620
 DB 1888 TCACTAGGGGAGACCAAGAGGTGGCTTAAATGACCACTGATCGTGTGATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGluValAlaIleThr 640
 DB 1948 AGTGTGCTTCATTTGACAGAGAAAGCGGAAAGTCCATATGGGTACAAAGTGTGGCCACA 2007
 QY 641 GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 DB 2008 GAAGTAGGGGGGTCTTCTTAAGCTCTGTGACAGATTCACCTGATCTTAATGATGTG 2067
 QY 661 AsnAspAsnProProArgLeuAlaValAspTyrThrGlyLeuPhePheCysHisProLeu 680
 DB 2068 AATGACAAACCTTCCAGGCTAGCCAGGACTTACACGGGCTTGTCTTCCCATCCCTCC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 DB 2128 AGTGCACCTGGAAAGTCTCATTTTCGAGGCTAGATGATGATCAACAATTATTCGGGGT 2187
 QY 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTyrGluValSerTyr 720
 DB 2188 CCCCATTTTACATTTTCCCTCGGCAAGTGAAGCTTACAAACACATCCGAAAGTTTCCAA 2247
 QY 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 DB 2248 ATCAATGGTACTCATCCGACACTGTCTACAGGCAACAGACTTTGAGGAGAGGGGTAT 2307
 QY 741 ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGluValIleValSerLeu 760
 DB 2308 GTGTCTTGATCCGACATCAATGATGGGGGTGGCCACCTTGAAGGCATTTGTTCTTAA 2367
 QY 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGlnThr 780
 DB 2368 CCAATTACATTCGACGTTGTGTGAGAGAAAGTTGTTCCGGCCAGACGATCCAGACT 2427
 QY 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 DB 2428 GGGATATCCCACTGTGGGCAATGGAGTGGTATCTCTGACCAACCTTCTGGGATGGT 2487
 QY 801 IleIleLeuAlaValValPheIleArgIleValValAspValGlyValAspAsnValGlu 820
 DB 2488 ATTAATTTTACAGTGTGTGTATCCGATTAAGAGAGATTAAGCAAGATTAATGTGAA 2547

QY 821 SerAlaGlnAlaSerGluValIleValProLeuArgSer 832
 DB 2548 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAGC 2583
 RESULT 2
 AAI29511
 ID AAI29511 standard; cDNA; 3345 BP.
 AC AAI29511;
 XX
 XX 12-OCT-2001 (first entry)
 DE Human intestinal peptide-associated transporter HPT-1 cDNA sequence.
 XX
 XX Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KM immunogenic; gene therapy; vaccine; colonic cancer; ss.
 OS Homo sapiens.
 XX
 XX WO200149716-A2.
 PN
 PD 12-JUL-2001.
 XX
 PF 29-DEC-2000; 2000MO-US35596.
 XX
 PR 30-DEC-1999; 99US-0476296.
 PR 10-JAN-2000; 2000US-0480321.
 PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.
 XX
 PA (COR-) CORIXA CORP.
 XX
 XX Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;
 PI King GE, Wang T, Jiang Y;
 DR WPI; 2001-441847/47.
 PT Colon tumor associated proteins and nucleic acids useful for the
 PT prevention, diagnosis and treatment of colonic cancer -
 XX
 XX Claim 2; Page 466-467; 472pp; English.
 CC The present invention describes colon tumour associated proteins (I) and
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and
 CC (II) may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate colon tumour associated protein (TCAP)
 CC expression, such as colonic cancer. For example, (I) and (II) may be
 CC used to treat disorders associated with decreased expression by
 CC rectifying mutations or deletions in a patient's genome that affect the
 CC activity of TCAPs by expressing inactive proteins or to supplement the
 CC patients own production of them. Additionally, (II) may be used to
 CC produce the TCAP proteins, by inserting the nucleic acids into a host
 CC cell culturing the cell to express the protein. (II) and its
 CC complementary sequences may also be used as DNA probes in diagnostic
 CC polymerase chain reaction (PCR) and hybridisation assays to detect and
 CC quantitate the presence of similar nucleic acids in samples, and
 CC therefore which patients may be in need of restorative therapy. (I) may
 CC also be used as antigens in the production of antibodies against TCAPs
 CC and in assays to identify modulators of TCAP expression and activity.
 CC Anti-(I) antibodies and antagonists may also be used to down regulate
 CC TCAP expression and activity. The anti-(I) antibodies may also be used
 CC as diagnostic agents for detecting the presence of TCAPs in samples
 CC (e.g. by enzyme linked immunosorbent assay (ELISA)). AAI28460 to AAI29512
 CC and AAI24494 to AAI24523 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.
 XX
 SQ Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Alignment Scores:

Pred. No.: 0
 Score: 4321.00
 Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 100.00%
 DB: 22

Length: 3345
 Matches: 832
 Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-10-025-380-1081 (1-832) x AA129511 (1-3345)

QY 1 MetIleuGlnAlaHisLeuHisSerLeuCysLeuMetLeuTyrLeuAlaThrGly 20
 DB ATATATCTTCAGGCGCATCTTCACTCCCTGCTCTTATGCTTTATTTGGCACTGGA 147
 QY 21 TyrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTyrGlu 40
 DB TATGGCCAAAGAGGGAAAGTTAGTGAACCCCTGAACCAATGACATTTCTATTATGA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheValAlaAsnProProAlaValThrPhe 60
 DB GGCCAAAGACCGAGTCAATTAATTCAGTTTAAGCCAAATCCTCTGCTGACCTTTT 267
 QY 61 GluLeuThrGlyGlnThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTyrTyr 80
 DB GAATTAATCTGGGAGAGACAGACATATTTGTGATAGAACGGAGGAGACTTCTGTATTA 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB AACAGACCTTGGACAGGAAACAAAGATCTACTACATCTCCAGTTGACGCCCTTGAC 387
 QY 101 AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn 120
 DB GCTAATGAATTAATAGAGAGGATCCAGTCCCTTACCATAGAGTGAAGAGACATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerLysTyrGlySerValAlaGlnAsnSerArg 140
 DB GACATATGACCCACCTTCTTCTCACTCAAGTACAGAGGCTCAGTAAGGACAACTTCGC 507
 QY 141 ProGlyLysProPheLeuTyrValAlaAsnAlaThrAspLeuAspProAlaThrProAsn 160
 DB CCAGGAAAGCCCTTCTGTATGTCATGACAGACCTGATGATCCGGCCACATCCCAAT 567
 QY 161 GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 DB GGCCAGCTTATATCCAGATTTGATCCAGCTTCCATGATCAACATGTCATATCTTT 627
 QY 181 GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn 200
 DB CAGATCAACAACAAAGGAGGACATCTCTTACCCGAGAGGATCTCAGAAATTGAT 687
 QY 201 ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlyGlnSer 220
 DB CTGCTTAAGAATCTTCTTATATCTGATGATCTCAGTGAAGGACATGGGAGGCAAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTyrLys 240
 DB GAGATTCCTTCAGTATACCAATCTGTGATATCATAGTACAGAAATATTTGGAAA 807
 QY 241 AlaProLysProValGluMetValGluAsnSerThrAspProHisproIleLysIleThr 260
 DB GCACCAAAACCTGTGAGATGAGTGAACCACTGATCTCCACCCCATCAAAATCACT 867
 QY 261 GlnValAlaGTPAsnAspProGlyAlaGlnTyrSerLeuValAspLysGluLysLeuPro 280
 DB CAGGTGGCGTGAATGATCCCGGTGACAAATATCTTAACTTGAACAAAGAAACCTGCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValThrGlnProLeuAspArg 300
 DB AGATTTCCATTTCAATTGACAGAGAGAGATATTTACGTGACTCAGCCCTTGGACGA 987
 QY 301 GluGluLysAspAlaTyrValPheTyrAlaValAlaLysAspGluTyrGlyLysProLeu 320
 DB GAAGAAAGAGATGATATGTTTATATGACGTGCAAGAGATGAGTACGAAACCACTT 1047

QY 321 SerTyrProLeuGluIleHisValLysValLysAspIleAsnAspAsnProThrCys 340
 DB TCTATCCCTGGAAATTTATGTAATAAGTAAATTAATTAATGATATCCACTCATATGT 1107
 QY 341 ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 DB CCGTCAACAGTAACCGTATTTGAGTCCAGAGAAATGAACGACTGGGTAAACGATACGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 DB ACCCTTACGTGACATGACAGAGATGAAGAAATACCTGCCAACAGTTTCTTAATCAAG 1227
 QY 381 IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 DB ATTGTGAGCAACTCCCAACTCCCAATGATGATCTTCTTATCCAAACCTATGCT 1287
 QY 401 GlyMetLeuGlnLeuAlaLysGlnSerLeuLysValGlnAspThrProGlnTyrAsnLeu 420
 DB GGAAATGTACAGTTACTTAAACAGTCTTGAAGAACCAAGATACTCTCAGTACAACTTA 1347
 QY 421 ThrIleGluValSerAspLysAspPheLysThrLeuCysPheValGlnIleAsnValIle 440
 DB ACGATAGAGGTGTCTGACAAAGATTCAGAACCCCTTGTGTGCAAAATCAACGTTAT 1407
 QY 441 AspIleAsnAspGlnIleProIlePheGluLysSerAspTyrGlyAsnLeuThrLeuAla 460
 DB GATATCAATGATATAGATCCCATCTTTGAAATATGATTAAGAAACCTGACTTGGCT 1467
 QY 461 GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro 480
 DB GAAGACAAACAACTTGGGTCCACATCTTAACCATCAGGCCACTATGCTGATGAGCCA 1527
 QY 481 PheThrGlySerSerLysIleLeuTyrHisIleIleValGlyAspSerGluArgLeu 500
 DB TTTTACTGGAGTCTAAATTTGTATCATATATGAAGGAGAACAGTGAAGGACGCTG 1587
 QY 501 GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAsp 520
 DB GGCGTTGACACAGATCCCATCCATCCACACCGGATATGCTATTAATAAAGCCTTGAT 1647
 QY 521 PheGluThrAlaAlaValSerAsnIleValPheValAlaGluAsnProGluProLeuVal 540
 DB TTTGAAACAGCAGCTGTTTCCATCAATGTTGTTCAAGACGAAATCTGAGCTCTATG 1707
 QY 541 PheGlyValLysTyrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal 560
 DB TTTGGGTGTAAGTACAAATGCAATGCTTTTGGCAAGTTCAGCTTATTTGACAGATGT 1767
 QY 561 AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla 580
 DB AATGAAACACTCAATTTTCCCAACAGTATTCCAAGCGAAATGCTGAGATGATGCT 1827
 QY 581 IleGlyThrLysValGlyAsnValThrAlaLysAspProGluLysAspIleSerTyr 600
 DB ATAGGACATTAAGTGGCAATGATGATGCTCCAAAGATTCAGAAAGTGTGACATAGCTAT 1887
 QY 601 SerLeuArgGlyAspThrArgGlyTyrPheLysIleAspHisValThrGlyGluIlePhe 620
 DB TCACTGAGGGGAAACCAAGAGGTGGCTTAAATTAATGACCAAGTACTGGTGAATCTTT 1947
 QY 621 SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaThr 640
 DB AGTGTGCTCATTTGACAGAGAGAGCCGGAAGTCCATATGGGTACAAAGTGGTGGCCACA 2007
 QY 641 GluValGlyLysSerLeuSerSerValSerGluPheHisLeuIleMetAspVal 660
 DB GAAGTAGGGGGTCTTCTTAAGCTCTGTCAAGATTCACCTGATCTTATAGGATGTG 2067
 QY 661 AsnAspAsnProProAlaGluAlaLysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 DB AATGACAAACCTCCAGGCTAGCCAGACTTACAGGGCTGTGTTCTTCTCCATCCCTTC 2127
 QY 681 SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700


```

Db      2128 AGTCACCTGGAAATCCATTTCGAGGCTAATATATATGATGACACTTATATTCGGGCT 2187
Qy      701  ProtiisPheThrPheSerLeuGlySerGlySerLeuGlnAenAptPGLuValSerLys 720
Db      2188 CCCATTTCATCTTTCCCTCGGAGTGGAAAGCTTACAAACGACTGGAGATTTCGAA 2247
Qy      721  IleAnGlyThRhiSaIaArgLeuSerThRArgHisThRAspPheGluArgAlaTyr 740
Db      2248 ATCAATGATGACTACATGCCGACTGTCTACCAAGCAGACAGACTTGAAGAGAGCGGAT 2307
Qy      741  ValValLeuIleAlaGlyIleAsnAspGlyValArgProProLeuGluGlyIleValSerLeu 760
Db      2308 GTGCTCTGTGATCCCATCAATGATGGGGGCTGGCCACCTTGAAAGCATTTGTTCTTTA 2367
Qy      761  ProValThRPhCySerSerCyValGluGlySerCySpPheArgProAlaGlyHisGlnThr 780
Db      2368 CCACTTCATCTTCGACTGTGTGTGGAGAGAGATTGTTCCGCGCAGAGGTCACCAAGCT 2427
Qy      781  GlyIleProThRValGlyMetAlaValGlyIleLeuLeuThRThRLeuLeuValIleGly 800
Db      2428 GGGATACCCACTGTGGGCGATGGCAGTTGGTATATCTGCTGACCACTTCTGTGATTGGT 2487
Qy      801  IleIleLeuAlaValValPheIleArgIleLysLysAspLysGlyLysAspAsnValGlu 820
Db      2488 ATATATTTTACAGATGTGTGTATTCGGATTAAGAGATTAAGCAAGATTAATGTTGAA 2547
Qy      821  SerAlaGlnAlaSerGluValLysProLeuArgSer 832
Db      2548 AGTCTCAAGCATCTGAAGTCACCTCTGAGAAAGC 2583

RESULT 3
ABN96791 standard; DNA; 3345 BP.
XX      ABN96791;
AC      13-AUG-2002 (first entry)
XX      DE      Gene #3289 used to diagnose liver cancer.
XX      KW      Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX      KM      metastatic liver tumour; cytostatic; expression profile; disease state;
XX      KW      disease progression; drug toxicity; drug efficacy; drug metabolism.
XX      OS      Homo sapiens.
XX      PN      WO200229103-A2.
XX      PD      11-APR-2002.
XX      PF      02-OCT-2001; 2001WO-US30589.
XX      PR      02-OCT-2000; 2000US-237054P.
XX      PA      (GENE-) GENE LOGIC INC.
XX      PI      Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX      DR      WPI; 2002-426119/45.
XX      PT      Diagnosing and detecting the progression of liver cancer,
XX      PT      hepatocellular carcinoma or metastatic liver tumor in a patient,
XX      PT      involves detecting the level of expression of two or more genes in a
XX      PT      liver tissue sample
XX      CI      claim 1; SEQ ID NO 3289; 298bp; English.
XX      CC      The invention relates to a novel method for diagnosing and detecting the
XX      CC      progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX      CC      tumor in a patient, and differentiating metastatic liver cancer from
XX      CC      hepatocellular carcinoma in a patient, involving detecting the level of
XX      CC      expression of two or more genes represented in ABN93503-ABN97455 in a

```

```

CC      tissue sample. The method of the invention has hepatotropic, and
CC      cytostatic activity. The method is useful for diagnosing and detecting
CC      the progression of liver cancer, hepatocellular carcinoma and metastatic
CC      liver carcinoma in a patient. The method is useful for identifying
CC      expression profiles which serve as useful diagnostic markers as well as
CC      markers that can be used to monitor disease states, disease progression,
CC      drug toxicity, drug efficacy and drug metabolism.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences.
XX      SQ      Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 3345
Score: 4321.00 Matches: 832
Percent Similarity: 100.008 Conservative: 0
Best Local Similarity: 100.008 Mismatches: 0
Query Match: 100.008 Indels: 0
DB: 24 Gaps: 0

US-10-025-380-1081 (1-832) x ABN96791 (1-3345)
Qy      1  MetIleuGlnAlaHisIleuHisSerLeuCyLeuLeuMetLeuTyrIleuAlaThrGly 20
Db      88  ATGATATCTTCAGGCCCATCTTCACTCCCTGTGCTTATATGCTTATATGGCAACTGGA 147
Qy      21  TyrGlyGluGluGluLysPheSerGlyProLeuLysProMetThRPhSerLysIleTyrGlu 40
Db      148  TATGGCCCAAGAGGGAGAAATTATGTGAGACCCCTAAACCCATVGCATTTTATATTAGAA 207
Qy      41  GlyGluGluProSerGlnIleIlePheGlnPheAlaAsnProProAlaValThrPhe 60
Db      208  GGCCCAAGAACCCGATCAATATATATTCAGTTTAAGCCCAATCTCTGCTGACTTTT 267
Qy      61  GlnLeuThRGLyGluThRAspAsnIlePheValIleGluArgGlyLeuLeuTyrTyr 80
Db      268  GAACATACTGGGAGACAGACACATATTTGTATGAAGACGGAGGACTTCTGTATTAC 327
Qy      81  AsnArgAlaLeuAspArgGluThRArgSerThRHisAsnLeuGlnValAlaAlaLeuAsp 100
Db      328  AACAGAGCCCTTGACAGGAGAAACAAATATATCAACATATTCAGATTGACAGCCCTGGAC 387
Qy      101  AlaAsnGlyIleIleValGluGlyProValProIleThRileGluValLysAspIleAsn 120
Db      388  GCTAATGGAAATTAATGTGAGGGGTCCAGTCCCTATACCACTAAGAGTGAAGGACATAC 447
Qy      121  AspAsnArgProThRPhLeuGlnSerLysTyrGluGlySerValAlaArgIleAsnSerArg 140
Db      448  GACAAATCGACCCAGCTTTCAGCTCAAGTCAAGAGGCTCACTAAGGAGCAAACTCTCGC 507
Qy      141  ProGlyLysProPheLeuTyrValAlaAsnAlaThRAspLeuAspAspProAlaThrProAsn 160
Db      508  CCAGAAAGGCCCTTCTGTGATGCAATGCCAGACCTGATATATCCGGCACTCCCAAT 567
Qy      161  GlyGlnLeuTyrTyrGlnIleValIleGlnLeuProMetIleAsnValMetTyrPhe 180
Db      568  GGCCAGCTTATATCAAGATGTGATCCAGCTTCCATGATCAACATGTCAATGTACTTT 627
Qy      181  GlnIleAsnAsnLysThRGLyAlaIleSerLeuThRArgGluGlySerGlnGluLeuAsn 200
Db      628  CAGATCAACAAACAAACGGAGCATCTCTTACCCGAGAGGATCTCGAATGTAAT 667
Qy      201  ProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMetGlyGlnSer 220
Db      688  CTGTCTAAGAAATCTTCTATATATCTGTGATCTCGTGAAGACATGGAGGCCAGAGT 747
Qy      221  GluAsnSerPheSerAspThRThRSerValAspIleIleValThRGLuGlnSerLys 240
Db      748  GAGAAATTCCTTCGATATACCATCTGTGATATATGTGACAGAAATATTTGGAA 807
Qy      241  AlaProLysProValGluMetValGluAsnSerThRAspProIleLysIleThr 260

```

Db GCACCAAAACCTGTGAGATGTGGAAAACCTCAACTGATCTCTCAACCCCATCAAAATCACT 867
 Qy GlnValArgTrrPaaAapProGluValAglIntYserLeuValAapLySGluLeuPro 280
 Db CAGGTGGCGTGAATGATCCGGTGCACAAAATTCCTTAAGTGAACAAGAAACCTGCACA 927
 Qy ArgPheProPheSerIleAspGlnGluGlyAspIleTyrValTrpGlnProLeuAspArg 300
 Db AGATTCCCATTTTCAATTGACCAAGAGAAATATTTCGTGACTGACCTGACCCCTTGACCGA 987
 Qy GluGluIysAspAlaTyrValPheTyrAlaValAlaIysAspGluTyrGlyLeuProLeu 320
 Db GAAGAAAAGATGATATGTTTTTATGCACTTGCAAGAGATGATGACGGAACCACTT 1047
 Qy SerTyrProLeuGluIleHisValIysValIysAspIleAsnAapAapProProThrCys 340
 Db TCATATCCGCTGGAATTCATGTAATAAGTTAAAGTATTAATGATTAATCCCACTCAACGT 1107
 Qy ProSerProValTrpValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly 360
 Db CCGTACCAAGTAAACCGTATTGAGGTCCAGAGAAATGACACCTGGTAAACGATTCGGG 1167
 Qy ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTyrArg 380
 Db ACCCTTACTGCACATGACAGGAGTGAAGAAATACTGCCACAGTTTCTTAACCTACAGG 1227
 Qy IleValGluGlnTrpProIysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 Db ATTGTGGAGCAAACTCCCAAACTTCCCATGATGACCTTCCTCAATCCAAACCTAATGCT 1287
 Qy GlyMetLeuGlnLeuAlaIysGlnSerLeuIysGlnAspTrpProGlnTyrAsnLeu 420
 Db GGAATCTTACGTTAGCTTAACAGTCTCTGAAGAGCAAGATACCTCCCTCACTACACTTA 1347
 Qy ThrIleGluValSerAspIysAspPheIysThrLeuCysPheValGlnIleAsnValIle 440
 Db ACGATAGAGGTGTCTGACAAAGATTCAAGACCTTTGTTTGTCGAATCAACGTATT 1407
 Qy AspIleAsnAspGlnIleProIlePheGluIysSerAspTyrGlyAsnLeuThrLeuAla 460
 Db GATATCAATGATCAATCCCATCTTGAATAATCAATGAAATCAAGAACTCACTTGTGCT 1467
 Qy GluAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaTrpAspAlaAspGluPro 480
 Db GAAGACCAAAACATTGGGTCCACCATCTTAACATCCAGGCCACGTGAGCTGAGAGGCA 1527
 Qy PheThrGlySerSerIleLeuTyrHisIleIleIysGlyAspSerGluGlyArgLeu 500
 Db TTTACTGGAGTCTTAATAATCTGATCATATCAATAAGGAGACAGTGAAGGAGCGCTG 1587
 Qy GlyValAspThrAspProHisThrAsnThrGlyTyrValIleIleIysLeuProLeuAsp 520
 Db GGGGTGACACAGATCCCATTAACCAACCGGATATGTCATTAATTAATAAACCTCTTGAT 1647
 Qy PheGluThrAlaAlaValSerAsnIleValPheIysAlaGluAspProGluProLeuVal 540
 Db TTTGAACACAGACGCTTTTCCACATCTGTCTCAAGACAGAAATCTCGACCTCTAGTG 1707
 Qy PheGlyValIysTyrAsnAlaSerSerPheAlaIysPheThrLeuIleValThrAspVal 560
 Db TTTGGTGTGAAGTCAATGCAAGTCTTTTGCCAACTTCAAGCTTATGTGACAGATGTG 1767
 Qy AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaIysValSerGluAspValAla 580
 Db AATGAAGACCTCAATTTTCCCAACACGATATTCAAAGGAAAGTCAGAGAGATGTACT 1827
 Qy IleGlyThrIysValGlyAsnValThrAlaIysAspProGluGlyLeuAspIleSerTyr 600
 Db ATAGGACCTTAAGTGGGCAATGTGACTGCCAAGATCCAGAAAGTCTTGACATTAAGCTAT 1887
 Qy SerLeuArgGlyAspThrArgGlyTrpLeuIysIleAspHisValThrGlyGluIlePhe 620
 Db TCAGTGAAGGAGACAAAGAGTGTGGCTTAATAATTGACCAAGTGAAGTGTGAGATCTTT 1947

Qy SerValAlaProLeuAspArgGluAlaGlySerProTyrArgValGlnValAlaIleThr 640
 Db AGTGTGCTTCATTTGACACAGAGAACCGGAGATTCATTCGGTACAAAGTGGGCCACA 2007
 Qy GluValGlyGlySerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal 660
 Db GAAGTGAAGGGGCTTCTTCACTGATGCTGATGACAGTTCACCTGATCTTATGATGTG 2067
 Qy AsnAspAsnProProAlaArgLeuAlaIysAspTyrThrGlyLeuPhePheCysHisProLeu 680
 Db AATGACAAACCTCCAGGCTAGCCAAAGACCTACACGAGCTGTGTCTTGCCATCCCTC 2127
 Qy SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly 700
 Db AGTGCACCTGGAAGTCTCATTTTTCGAGGCTACTGATGATGATGATGACATTAATTCGGGCT 2187
 Qy ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluValSerIys 720
 Db CCCCATTTTACATTTTCCCTCGGACATGGAACTTCAAAAACGACTGGGAACTTCCAAA 2247
 Qy IleAsnGlyThrHisAlaIleArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
 Db ATCAATGTGATCTATGCTGCGACATGCTTACACAGGACACACACTTGAAGAGAGCGGTAT 2307
 Qy ValValLeuIleArgIleAsnAspGlyGlyArgProProLeuGlnGlyIleValSerLeu 760
 Db GTGCTCTTGAATCCGATCAATGATGAGGGGTGGCCACCTTGGAAGGATGTTCTTTA 2367
 Qy ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlnHisGlnThr 780
 Db CCAAGTACATTTGCAAGTGTGTGGAAGAAAGTTGTTCCGGCCAGAGAGCTCACAGACT 2427
 Qy GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db GGGATACCCCACTGGGCAAGGCAAGTGTGATCTGCTGACCACTTGTGATGATGTG 2487
 Qy IleIleLeuAlaValAlaPheIleArgIleIysLeuAspIysGlyIysAspAsnValGlu 820
 Db ATTAATTTTACAGTGTGTGTTATCCGCTAAAGAGATTAAGCAAGTAATGTTGA 2547
 Qy SerAlaGlnAlaSerGluValIysProLeuArgSer 832
 Db AGTGTCAAGCATCTGAAGTCAAACTCTGAGAAGC 2583
 Db 2548 AGTGTCAAGCATCTGAAGTCAAACTCTGAGAAGC 2583
 RESULT 4
 ABL62444
 ID ABL62444 standard; DNA: 3345 BP.
 AC ABL62444;
 XX
 DT 15-MAY-2002 (first entry)
 XX
 DE Colon adenocarcinoma related gene sequence SEQ ID NO:781.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytosolic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN M02001.94629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.

PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
 PI Soppet DR, Weaver Z;
 XX
 DR WPI; 2002-188264/24.

XX Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set

XX Claim 1; SEQ ID 781; 44pp; English.

XX The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cyclostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

Alignment Scores:

Pred. No.: 0 Length: 3345

Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-025-380-1081 (1-832) x ABL62444 (1-3345)

QY 1 MetIleuGlnAlaHisIleuHISserIleuCYaleuLeuMetIleuAlaThrGly 20
 DB 88 ATGATCTTCAAGCCCATCTTCACTCCCTGTGTCTTATGCTTATTTGGCACTGGA 147
 QY 21 TyrGIyGlnGluGlyPheSerGIyProLeuLYsProMetThrPheSerIleYrGlu 40
 DB 148 TATGGCAAGAGGGGAAGTTTATGTGACCCCTGAACCATGACATTTCTATTATGA 207
 QY 41 GlyGlnGluProSerGlnIleIlePheGlnPheLYsAlaAsnProAlaValThrPhe 60
 DB 208 GGCCAAAGAACCGAGTCAAAATTATATTCAGTTTAAAGGCCAATCTCTGCTGACTTTT 267
 QY 61 GluIleuThrGIyGluThrAspAsnIlePheValIleGluArgGluGlyLeuIleuYrYr 80
 DB 268 GAACCTAACCTGGGAGAGACACACATATTTGTATGAAAGGAGGACTTGTATATAC 327
 QY 81 AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB 328 AACAGAGCCTTGAACAGGGAACACATCTACTCAAACTCCAGTTGAGCCCTGGAC 387
 QY 101 AlaAsnGlyrIleIleValGluGlyProValProIleThrIleGluValLYsAspIleAsn 120
 DB 388 GCTAATGGAATTAATGAGAGGGTCCAGTCCCTATACCATAGAGTGAAGGACATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerIleYrGIyGlySerValArgGlnAsnSerArg 140
 DB 448 GACAAATGACCCAGCTTCTCCAGTCAAGTAAAGTGAAGGCTCACTAAGCGCAACTCTCCG 507
 QY 141 ProGIyLYsProPheLeuYrValAsnAlaThrAspLeuAspProAlaThrProAsn 160
 DB 508 CCAGAAAGCCCTCTCTGTGTATGTCATGCCACAGACTGATGATCCGCCATCTCCAA 567
 QY 161 GlyGlnLeuYrYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetYrPhe 180
 DB 568 GGCCAGCTTATATACAGATGTCATCCAGCTTCCCATGATCAACATGTCATGACTTT 627
 QY 181 GlnIleAsnAsnLYsThrGlyAlaIleSerLeuThrArgGluGlySerGlnIleuAsn 200
 DB 628 CAGATCAACAAACAAACGGAGGCACTCTCTTAAACCGAGAGGATCTCAGGAATTGAAT 687
 QY 201 ProAlaLYsAsnProSerTYrAsnLeuValIleSerValLYsAspMetGIyGlnSer 220
 DB 688 CTGTCTAAGATCTCTCTTAATCTGTGATCTCGTAAGAGACATGGAGGCCAGAGT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTrpLYs 240
 DB 748 GAGAAATCTCTTCAAGTATACCATCTGTCATATCATGTGACAGAAATATTGGAAA 807
 QY 241 AlaProLYsProValGluMetValGluAsnSerThrAspProHisProIleLYsIleThr 260
 DB 808 GAACCAAAACCTGTGAGATGTGGAAACTCAACTGATCCCTCACCCATCAAAATCACT 867
 QY 261 GlnValArgTrpAsnAspProGlyAlaGlnYrSerLeuValAspLYsGluLYsLeuPro 280
 DB 868 CAGGTCCGTGATATCATCCGGTGCACAAATTTCTTATGTTGACAAAGAGAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGluLYsAspIleYrValThrGlnProLeuAspArg 300
 DB 928 AGATTCCATTTTCAATTATACACAGAGAGAGATTTTACGTACACAGCCCTTGACCGA 987
 QY 301 GluGluLYsAspAlaLYsValPheYrValAlaValAlaLYsAspGluYrGIyLYsProLeu 320
 DB 988 GAAGAAAGAGATGATATGTTTATATGACATGTAAGAGATGAGTGAAGAAACCACTT 1047
 QY 321 SerTYrProLeuGluIleHisValLYsValLYsAspIleAsnAspAsnProProThrCys 340

PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX

DR WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set

XX Claim 1; SEQ ID 2761; 44pp; English.

CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophagel, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 3345 BP; 951 A; 746 C; 720 G; 928 T; 0 other;

XX Alignment Scores:

Pred. No.: 0 Length: 3345
 Score: 4321.00 Matches: 832
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-10-025-380-1081 (1-832) x ABL64424 (1-3345)

QY 1 MetllleuGlulAlahlsleuHlssetleuCyaleuMeuMeuUryleuAlatnrgly 20
 DB 88 ATGATGCTTCAGGCGCATCTTCACTCCCTGTGCTTATGCTTATTATTTGGCACTGGA 147
 QY 21 TTTGGLGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
 DB 148 TATGTCACAGAGGGGAGATTAGTGGACCCCTGAAACCCATGACATTTCTATTATTTGAA 207

QY 41 GlyGlnGlnProSerGlnIlellepheGlnPheValaasnProProAlaValThrPhe 60
 DB 208 GGCACAGAACCGAGTAAATTAATTCAGCTTTAAAGCCAACTCTCTGTCGACTTTT 267
 QY 61 GlnLeuThrGlyGlnUthrAspAsnIlelleValIleGlnUrgGlnGlyLeuLeuTyr 80
 DB 268 GAACCTAAGTGGGGAGACAGAACATATTTTGGATGTAAGACGGGAGGAGCACTTGTATTAC 327
 QY 81 AsnArgAlaLeuAspArgGlnUthrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 DB 328 AACAGAGCCTTGACAGAGGAAACAAAGATCTACTCAAACTCCAGGTTCCAGCCCTGGAC 387
 QY 101 AlaAsnGlyIlelleValGlnGlyProValProIleThrIleGlnValIysAspIleAsn 120
 DB 388 GCTAATGGAATTAATGATGGAGGTCAGTCCCTATACCAATGAAAGGATCAAC 447
 QY 121 AspAsnArgProThrPheLeuGlnSerIysTyrGlnGlySerValArgGlnAsnSerArg 140
 DB 448 GACATCGAACCCAGCTTTCTCCAGTCAAGTACGAAGGCTCAGTAAAGCACTCTCGG 507
 QY 141 ProGlyIlyProPheLeuTyrValaAsnAlaThrAspLeuAspProAlaThrProAsn 160
 DB 508 CCAAGAAACCCCTTCTTGTATGTCATAGCACAGACCTGATGATCCGCCACTCCCAAT 567
 QY 161 GlnLeuTyrTyrGlnIlelleValIleGlnLeuProMetIleAsnAsnValMetTyrPhe 180
 DB 568 GGCACGCTTATTAACAGATGTCATCCAGCTTCCATATACAAATGTCATGTCCTT 627
 QY 181 GlnIleAsnAsnIysThrGlyAlaIleSerLeuThrArgGlnGlySerGlnGlnLeuAsn 200
 DB 628 CAGATCAACAAACAAACGGAGAGCATCTCTTACCAGAGAGGATCTCAGGAATTGAAT 687
 QY 201 ProAlaIlyAsnProSerTyrAsnLeuValIleSerValIysAspMetGlyGlnSer 220
 DB 688 CCGTCMAAATCCCTTCTTAATCTGATGATCTCAAGTAAAGACATGGAGGCCAGAT 747
 QY 221 GluAsnSerPheSerAspThrThrSerValaAspIlelleValThrGlnAsnIleTyrPlys 240
 DB 748 GAGAAATCTTCAAGTATACCACTGTCGATATCATATGACAGAAATTTTGGAAA 807
 QY 241 AlaProIysProValGlnUthrValGlnAsnSerThrAspProHisProIleIysIleThr 260
 DB 808 GACCCAAACCTGTGGAGATGGGAAACCTCAAGTACCTCACCCCAACCAATCACT 867
 QY 261 GlnValaIysTyrAsnAspProGlyAlaGlnTyrSerLeuValAspIlyGlnUthrPro 280
 DB 868 CAGGTCCGTGGAATGATCCCGTGCACAAATTCCTTATGTTGACAAAGAAAGCTGCCA 927
 QY 281 ArgPheProPheSerIleAspGlnGlnGlnIysAspIleTyrValThrGlnProLeuAspArg 300
 DB 928 AGATTCCTCAATTTCAATTAACAGAGAGAAATTTAGTACCTCAGCCCTTGGACCA 987
 QY 301 GlnGlnUlyAspAlaTyrValPheTyrAlaValAlaIysAspGlnUtyrGlyUthrProLeu 320
 DB 988 GAAAGAAAGAGATGATATTTTATGAGTTGCAAGATGAGATGAGAAACCACTT 1047
 QY 321 SerTyrProLeuGlnIleHisValIysValIysAspIleAsnAspAsnProThrTyrCys 340
 DB 1048 TCATATCCCTCGAATTCATGTAAGTAAAGTAAATATATATATATATATATATATAT 1107
 QY 341 ProSerProValThrValPheGlnValGlnGlnGlnUthrGlnUthrGlnUthrIleGly 360
 DB 1108 CCGTCAACGATTAACGATTTTGAAGTCCAGGAGAAATGAACGACTGGTAAACGATTCGGG 1167
 QY 361 ThrLeuThrAlaHisAspArgAspGlnGlnUthrAlaAsnSerPheLeuAsnTyrArg 380
 DB 1168 ACCCTTACTGACATGACAGAGATGAAGAAATATATGCAACAGTTTCTTAATCAACAGG 1227
 QY 381 IleValGlnGlnThrProIysLeuProMetAspGlyLeuPheLeuIleGlnThrTyrAla 400
 DB 1228 ATTTGAGACAAATCCCAACTCTCCATGTAGACTCTTCTAATCAACCAATATGCT 1287

QY	401	GLYVETLEUGNLEUVALYVSGINSEIRLEUYSIYSGINAPPTPRGGINTYRASP	420
Db	1288	GGAAATGTACAGTTAGCTAAACAGCTCTTGAAAGACAGATACCTCCAGTACA	1347
QY	421	ThrlleGuValSerAspLysAspPheLysThrLeuCySpheValGlnleAsnValIle	440
Db	1348	ACGATAGAGGCTGTGCACAAAGATTTCAGAACCTTTGTGTTGTGCAAAATCA	1407
QY	441	AspIleAsnAspGlnIleProIlePheGluYsSerAspTYRGIYAsnLeuThrLeuVal	460
Db	1408	GATATTCATGATCATCGATCCCATCTTTGAAAACACAGATTGTGAAACCTGAC	1467
QY	461	GluAspThrAsnIleGlySerThrIleLeuThrIleGlnIleAspAlaAspGluPro	480
Db	1468	GAAAGCACAAACATTTGGGTCCACCATCTTAACATCCAGGCCACTGATCTGATGAC	1527
QY	481	PheThrGlySerSerLysIleLeuTYRHisIleIleLysGlyAspSerGluGlyArgLeu	500
Db	1528	TTTATCTGGGAATTCCTAAATCTGTATCATATCAATAAGGAGACAGTGAAGAG	1587
QY	501	GlyValAspThrAspProHisThrAsnThrGlyTYRValIleIleLysLysProLeuAsp	520
Db	1588	GGGGTTCACACAGATCCCCATACCAACCCGGATATGCTAATAATTAACACCTTG	1647
QY	521	PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAspProGluProLeuVal	540
Db	1648	TTTGGAAACAGAGCGTTTCCAACTTGTTCCAAAGCAAAATCCAGAGCTCTTA	1707
QY	541	PheGlyValLysTYRAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal	560
Db	1708	TTTGGTGTGAAGTACATACGACAGTCTTTGGCAAGTTCAGGCTTATGTGACAGATG	1767
QY	561	AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla	580
Db	1768	AATGACAGCACTCAATTTTCCCAACAGCTATTCCAAAGCAAAAGTCAGAGAGATG	1827
QY	581	IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTYR	600
Db	1828	ATAGGACATTAAGGTGGGACATGTGACTGCCAAGATCCAAAGAGCTTGACATTA	1887
QY	601	SerLeuAspGlyAspThrArgGlyTYRLeuLysIleAspHisValThrGlyGluIlePhe	620
Db	1888	TCACTGAGGGGAGACACAGAGGTGGCTTAATAATTGACCAAGTACCTGTGAGATCT	1947
QY	621	SerValAlaProLeuAspArgGluAlaGlySerProTYRArgValGlnValAlaThr	640
Db	1948	AGTGTGGCTCCATTTGGACAGAGAACCCGGAAGTCCATATCGGGTCAAGTGGTGC	2007
QY	641	GluValAlaGlyLysSerSerLeuSerSerValSerGluPheHisIleuIleLeuMetAspVal	660
Db	2008	GAAAGTAGGGGGGTCTTCTTAAGCTCTGTGTCAAGATTCACCTGATCTTAATG	2067
QY	661	AsnAspAsnProPoxArgLeuAlaLysAspTYRThrGlyLeuPhePheCyHisProLeu	680
Db	2068	AATGACAAACCTTCCAGGCTACCAAGACATACAGGGCTTGTCTTCCCATTCCTC	2127
QY	681	SerAlaProGlySerLeuIlePheGluAlaThrAspAspAspGlnHisLeuPheArgGly	700
Db	2128	AGTGAACCTGGAAAGTCTCATTTTCCAGAGCTATCTGATGATCAGACACTTA	2187
QY	701	ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTYRGIYAsnLys	720
Db	2188	CCCCATTTTACATTTTCCCTCCGACAGTGGAACTTAACAAACGACTGGGAATTTT	2247
QY	721	IleAsnGlyThrHisAlaArgLeuSerThrArgHisIleThrAspPheGluGluArgAlaTYR	740
Db	2248	AATCAATGAGTACTAGCCCGAGCTGTCTACAGAGCACACAGACTTTGAGGACAGG	2307
QY	741	ValIleValLeuIleArgIleAsnAspGlyGlyArgProProLeuGluGlyIleValSerLeu	760
Db	2308	GTCGCTTGATCCGCATCATATATGGGGGTCCGCCACCTTGGAAAGCATTTCTTT	2367
QY	761	ProValThrPheCySerCyValGluGlySerCySpheArgProAlaGlyHisIleThr	780

Db	2368	CCAGTTACATCTTCGACAGTTGTGGAAAGGAAGTTGTTCCGGCAGACGCTCACCAAGCT	2427
Qy	781	GLYIleProthrValGIyMetaIaValGIyIleLeuLeuThrIleLeuValIleGIy	800
Db	2428	GGGATACCCCACTGTGGCAGTGGAGTGGTATCTGTACCAACCCCTTCGTGGATTGCT	2487
Qy	801	IleIleLeuIaValValPheIleArGIleLyAsApLySGIyLyAsAaenValGIy	820
Db	2488	ATATATTTAGCAGTTGGTGTTCATCCCATTAAGAAGATTAAGCAAAAGTATATGTTGAA	2547
Qy	821	SeAlaGIaIaSeGIuValLyAsProLeuAsySer	832
Db	2548	AGTGCTCAAGCATCTGAACTCAAACTCTGAGAAGC	2583
RESULT	6		
ID	ABN97312	standard; DNA; 3654 BP.	
AC	ABN97312;		
XX	13-AUG-2002	(first entry)	
DE	Gene #3810	used to diagnose liver cancer.	
XX	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;		
KM	metastatic liver tumour; cytostatic; expression profile; disease state;		
KM	disease progression; drug toxicity; drug efficacy; drug metabolism.		
XX	Homo sapiens.		
XX	MO200229103-A2.		
PN	11-APR-2002.		
XX	02-OCT-2001; 2001MO-US30589.		
PF	02-OCT-2000; 2000US-237054P.		
PR	(GENE-) GENE LOGIC INC.		
XX	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;		
PI	WPI; 2002-426119/45.		
XX	Diagnosing and detecting the progression of liver cancer,		
PT	hepatocellular carcinoma or metastatic liver tumor in a patient,		
PT	involves detecting the level of expression of two or more genes in a		
PT	liver tissue sample		
XX	Claim 1; SEQ ID NO 3810; 298bp; English.		
PS	The invention relates to a novel method for diagnosing and detecting the		
XX	progression of liver cancer, hepatocellular carcinoma or metastatic liver		
CC	tumour in a patient, and differentiating metastatic liver cancer from		
CC	hepatocellular carcinoma in a patient, involving detecting the level of		
CC	expression of two or more genes represented in ABN93503-ABN97455 in a		
CC	tissue sample. The method of the invention has hepatotropic, and		
CC	cytostatic activity. The method is useful for diagnosing and detecting		
CC	the progression of liver cancer, hepatocellular carcinoma and metastatic		
CC	liver carcinoma in a patient. The method is useful for identifying		
CC	expression profiles which serve as useful diagnostic markers as well as		
CC	markers that can be used to monitor disease states, disease progression,		
CC	drug toxicity, drug efficacy and drug metabolism.		
CC	Note: The sequence data for this patent did not form part of the printed		
CC	specification, but was obtained in electronic format directly from WIPO		
CC	at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;		
Alignment Scores:	0	Length: 3654	
Pred. No.:	4303.00	Matches: 828	

Percent Similarity:	99.76%	Conservative:	2
Best Local Similarity:	99.52%	Mismatches:	0
Query Match:	99.58%	Indels:	2
DB:	24	Gaps:	0
US-10-025-380-1081 (1-832) x ABN97312 (1-3654)			
QY	1	MettIeuglnalAhIaIeUuHIsSeSerLeuCYsLeuUmeLeuTYrLeuAlaThrGly	20
DB	93	ATGATACCTCAGGCCCACTCTTCACTCCCTGTCGTCTTATGCTTATTTGGCACTACGA	152
QY	21	TYrGlyGlnGluGlyLysPheSerGlyProLeuLysProMetThrPheSerIleTYrGlu	40
DB	153	TATGGCCAAAGGGGGAAAGTTTATGTGAGACCCCTGAACCCCATGACATTTTCAATTATGAA	212
QY	41	GLyGlnGluProSerGlnIleIlePheGlnPheValIaAsnProProlAvalThrPhe	60
DB	213	GGCCAAACCCGATCAATTAATTAATTCAGTTTAAGGCCAATCCTCGCTGTGACTTTT	272
QY	61	GLuLeuThrGlyGluThrAspAsnIlePheValIleGluArgGluGlyLeuLeuTYrTYr	80
DB	273	GAACCTAACCTGGGAGACAGACACATATTTGTGATAGACGGGAGGACTTCTGTATTAAC	332
QY	81	AsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp	100
DB	333	AACAGAGCCTTGGACAGGGAAACAAAGATCTACTACAACTCCAGGTTGACGCCCTGGAC	392
QY	101	AlaAsnGlyIleIleValGluGlyProValProIleThrIleGluValLysAspIleAsn	120
DB	393	GCTAAATGGAATTAATAGAGAGGCTCCAGCTCCATACCAATAAAGTAAAGGAGACATCAAC	452
QY	121	AspAsnArgProThrPheLeuGlnSerLysTYrGlyGlySerValArgGlnAsnSerArg	140
DB	453	GACATGAGACCCAGCTTCTCCAGTCAAGTACAAAGGCTCAAGTAAGGCAACTCTCGC	512
QY	141	ProGlyLysProPheLeuTYrValAsnAlaThrAspLeuAspProAlaThrProAsn	160
DB	513	CCAGGAAAGCCCTTCTGTATGTCAATGCCACAGACCTGATGATCGGGCACTCCCAAT	572
QY	161	GLyGlnLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsnValMetTYrPhe	180
DB	573	GGCCAGCTTTAATTAACGATGTCATCCACTTCCCAATGATCAACAATGTCATGATCTTT	632
QY	181	GlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySerGlnGluLeuAsn	200
DB	633	CAGATCAACAACAAAGGAGGCACTCTCTTAACCGAAGGGATCTCAGGAATTAAT	652
QY	201	ProAlaLysAsnProSerTYrAsnLeuValIleSerValLysAspMetGlyGlnSer	220
DB	693	CTGCTAAGAACTCTTCTTAATCTGATCTCAGTGAAGGACATGGAGGCGCAGAGT	752
QY	221	GLuAsnSerPheSerAspThrThrSerValAspIleIleValThrGluAsnIleTrpLys	240
DB	753	GAAATATCTCTTCAAGTATACCAATCTGTGATATATATGACAGGAATATTTTGAAA	812
QY	241	AlaProLysProValGluMetValGluAsnSerThrAspProHisProIleLysIleThr	260
DB	813	GCACCAAAACCTTGAGATGGATGGGAAAACCTCAACTGATCTCTCAACCCCATCAAAATCACT	872
QY	261	GlnValArgTrpAsnAspProGlyAlaGlnTYrSerLeuValAspLysGluLysLeuPro	280
DB	873	CAGGTCCGCTGAATGATCCCGGTGCACATATTCCTAGTTGCAAAAGAGAGCTGCCA	932
QY	281	ArgPheProPheSerIleAspGlnGluGlyAspIleTYrValThrGlnProLeuAspArg	300
DB	933	AGATTCACATTTTCAATGACAGGAAAGGATATTTACGTGACTCAGCCCTTGAGCCGA	992
QY	301	GluGluLysAspAlaTYrValPheTYrAlaValAlaLysAspGluTYrGlyLysProLeu	320
DB	993	GAAGAAAAGATGCAATATGTTTTTATGCAAGTTGCANAAGATGAGTACGAAAACCACTT	1052
QY	321	SerTYrProLeuGlnIleHisValLysValLysAspIleAsnAspAsnProThrTYrS	340

DB	1053	TCATATCCGCTGGAAATTCATGTAAGTAAAGATATTAATGATATCAACCTACATGT	1112
QY	341	ProSerProValThrValPheGluValGlnGluAsnGluArgLeuGlyAsnSerIleGly	360
DB	1113	CCGTCAACAGTAACCGTATTTGAGTCCAGGAATGAACGATGGGTAAACAGATTCGGG	1172
QY	361	ThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPheLeuAsnTYrArg	380
DB	1173	ACCTTACTGCACATGACAGGATGAGAAAATACGTGCCAAGCTTTTCTAACTACAGG	1232
QY	381	IleValGluGlnThrProLysLeuProMetAspGlyLeuPheLeuIleGlnThrTYrAla	400
DB	1233	ATTGTGAGCAAACTCCAAACTCCCATGAGTGGACTCTTCCTAATCCAAACTGATGTCT	1292
QY	401	GLMetLeuGlnLeuAlaLysGlnSerLeuLysGluAspThrProGlnTYrAsnLeu	420
DB	1293	GGATGTATAGTATCAATCAACGTCCTTGAAAGACAGATATCTCTCGTCAACTTA	1352
QY	421	ThrIleGluValSerAspLysAspPheLysThrLeuCYsPheValGlnIleAsnValIle	440
DB	1353	ACGATAGAGGTGCTGACAAAGATTTCAAGACCTTTGTTTGTGCAAATCAAGCTTATT	1412
QY	441	AspIleAsnAspGlnIleProIlePheGluLysSerAspTYrGlyAsnLeuThrLeuAla	460
DB	1413	GATATCAATGATCGAACCCCATCTTGAATAATCAGATTATGAAAACCTGACTTGTGCT	1472
QY	461	GLuAspThrAsnIleGlySerThrIleLeuThrIleGlnAlaThrAspAlaAspGluPro	480
DB	1473	GAAGACAAACAATGTGGTCCACATCTTAACATCCAGGCCCTAGTGTATGAGCCA	1532
QY	481	PheThrGlySerSerLysIleLeuTYrHisIleIleLysGlyAspSerGluGlyArgLeu	500
DB	1533	TTTACTGGAGTTCTTAATTTCTGTATCATATATAAAGGAGACAGTGAAGGACCCCTG	1592
QY	501	GlyValAspThrAspProHisThrAsnThrGlyTYrValIleIleLysLysProLeuAsp	520
DB	1593	GGGGTTGACACAGATCCCATCAACACCGGATATGTCAATATTAATAAAGCCTCTTGAT	1652
QY	521	PheGluThrAlaAlaValSerAsnIleValPheLysAlaGluAsnProGluProLeuVal	540
DB	1653	TTTGAACACGACACTTTTCCACATGTGTTCAAGAGAAATCTGAGCTCTAGTG	1712
QY	541	PheGlyValLysTYrAsnAlaSerSerPheAlaLysPheThrLeuIleValThrAspVal	560
DB	1713	TTTGTGTGAAGTACATGCACTTTTGGCCAGTTCAAGCTTATTTGTGACAGATGTG	1772
QY	561	AsnGluAlaProGlnPheSerGlnHisValPheGlnAlaLysValSerGluAspValAla	580
DB	1773	AATGAAGCACCTCAATTTTCCCAACGATATTCAGAGCGAAAGTCAGTGAGATGAGCT	1832
QY	581	IleGlyThrLysValGlyAsnValThrAlaLysAspProGluGlyLeuAspIleSerTYr	600
DB	1833	ATAGGACATTAAGTGGCAATGATGATGCCAAGAGTCCAGAAAGTGTGACATTAACCTAT	1892
QY	601	SerLeuArgGlyAspThrArgGlyTrpLeuLysIleAspHisValThrGlyGluIlePhe	620
DB	1893	TCACTAGGGGAGACACAAAGGTTGGCTTAATAATTGACCAAGTCACTGGTGAATCTTT	1952
QY	621	SerValAlaProLeuAspArgGluAlaGlySerProTYrArgValGlnValAlaThr	640
DB	1953	AGTGTGGCTCCATTGGACAGAGAAGCGGAAGTCCATATCGGGTACAAGTGGGCGACA	2012
QY	641	GluValGlyLysSerSerLeuSerSerValSerGluPheHisLeuIleLeuMetAspVal	660
DB	2013	GAAGTAGGGGGGCTTCTTGAAGCTTGTGTGAGATTCACCTGATCTTATGAGATGG	2072
QY	661	AsnAspAsnProProArgLeuAlaLysAspTYrThrGlyLeuPhePheCYsHisProLeu	680
DB	2073	AATGACAACTCTCCAGGCTAGCCAAAGACTACACGGGCTGTCTTCTTCCATCCCTTC	2132
QY	681	SerAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHisLeuPheArgGly	700
DB	2133	AGTGCACTGGAAGTCAATTTTTCGAGGCTACGATGATGATCAGACATTATTTGGGGGT	2192

```

Qy 701 ProHisPheThrPheSerLeuGlySerGlySerLeuGlnAsnAspTrpGluValSerLeu 720
Db 2193 CCCCATTTTACATTTTCCCTCGGAGGAAAGCTTACAAAACGACTGGAAAGTTTCCAAA 2252
Qy 721 IleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPheGluGluArgAlaTyr 740
Db 2253 ATCATAGTACTACTATGCCGACTGTCTACACGACACAGAGTTTGAAGAGAGGAGTAT 2312
Qy 741 ValValLeuIleArgIleAsnAspGlyGlyArgProPheLeuGluGlyIleValSerLeu 760
Db 2313 GTCGCTGTGATCCCATCAACAGATGGGGGTGCGCACCCCTTGAAGGCATTTGTTCTTTA 2372
Qy 761 ProValThrPheCysSerCysValGluGlySerCysPheArgProAlaGlyHisGluThr 780
Db 2373 CCACTTACATTTCTGCAAGTTGTGTGAAGAAAGTTGTTCCGGACACAGAGTACCACAACT 2432
Qy 781 GlyIleProThrValGlyMetAlaValGlyIleLeuLeuThrThrLeuValIleGly 800
Db 2433 GGGATACCCACTGTGGGCAATGGCAGTTGGTATACGTGGACCACTTCTGTGATTGGT 2492
Qy 801 IleIleLeuAlaValIleValPheIleArgIleValLeuAspGlyGlyValAspAsnValGlu 820
Db 2493 ATATATTTAGCAGTGTGTTTATCCGATAAAGAGTAAAGCAAAAGATTAATGTTGAA 2552
Qy 821 SerAlaGlnAlaSerGluValIleValProLeuArgSer 832
Db 2553 AGTGTCTCAGCATCTGAAGTAAACCTCTGAAAGC 2588

RESULT 7
ABL62484
ID ABL62484 standard; DNA; 3654 BP.
AC ABL62484;
XX
XX 15-MAY-2002 (first entry)
DT
XX
DE Colon adenocarcinoma related gene sequence SEQ ID NO:821.
XX
KM Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KM cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KM gene; ds.
XX
OS Homo sapiens.
XX
PN W0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
XX 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.

```

```

PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX
XX Claim 1, SEQ ID 821; 44pp: English.
XX
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytoskeletal
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX
XX Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 0 Length: 3654
Score: 4303.00 Matches: 828
Percent Similarity: 99.76% Conservative: 2
Best Local Similarity: 99.52% Mismatches: 0
Query Match: 99.52% Indels: 0
DB: 24 Gaps: 0

US-10-025-380-1081 (1-832) x ABL62484 (1-3654)
Qy 1 MetIleLeuGlnAlaHisLeuHisSerLeuCysLeuLeuMetLeuTyrLeuAlaThrGly 20
Db 93 ATGATACTTACGCGCCATCTTCACTCCCTGTGCTTCTTATATTTGGCACTGCA 152
Qy 21 TyrGlyGlnGluGlyValPheSerGlyProLeuLeuPheMetThrPheSerIleTyrGlu 40
Db 153 TATGCGCAAGAGGGGAAGTTTACTGAGACCCCTGAACCCATGACATTTCTATTATGAA 212

```

QY 41 GLYGLNGLUPROSERGINILEPHEGLNPHELYSLAASNPPOVALAVALTHRPH 60
DB 213 GGGCCAAAGACCGATCAATTAATATTCAGTTTAAGCCAAATCCTCGCTGACATTTT 272
QY 61 GNLNLEUTHRGVGLNTHRRASPAENILEPHEVALILEGLNLRGGLUGLYLEULEUTHYTYR 80
DB 273 GAACTTAACGTGGGAGACAGACACATATTTGTGATACACGGAGGAGCTTCGTGATTAAC 332
QY 81 ASNAHSLALEUASPARGLNTHRRASERETHRHSASNPUGNVALAIAALEUASP 100
DB 333 AACAGAGCCTTGACAGAGAAACAAGATCTACTACAAATCTCCAGTTGAGACCCCTGAC 392
QY 101 ALASNGLYILEILEVALIGLUGLYPROVALPROILETHRIEGLNVALYASPILEASN 120
DB 393 GCTAAAGSATTATAGGAGGGGCTCAGCTCCATACCATMAAAGTAAGGAGCATCAAC 452
QY 121 ASPASNPPOPHRPHLEUGNLSERLEYGLNGLYSERVALARGGLNANSEARAG 140
DB 453 GACATTCGACCCAGTTCTCCAGTCAAGTACGAAGGCTCAGTAAGGCAAACTCTGCG 512
QY 141 PROGLYASNPPOHLEUTHYRVALASNALETHRASPENASNPPOVALATHRPROASN 160
DB 513 CAGAGAAAGCCCTCTTGATGTCANCCAGACAGACCTGATGATCGGCCATCTCCAAAT 572
QY 161 GLYGLNLEUTHYTYRGLNILEVALILEGLNLEUPROKETILEASNPVALMECTYRPH 180
DB 573 GGCACACTTATTAACAGATTGCATCCAGCTCCATGATCAACATGTCATGTACTTT 632
QY 181 GNLILEASNPANLYSTHRGLYALAIIESERLEUTHRARGGLUGLYSERGLNGLULEASN 200
DB 633 CAGATCAACAAACAAACGAGGAGCATCTCTTAACCGAGAGGAGTCTCGAATTAAGAT 692
QY 201 PROALALYASNPPOSERLEYRASNLEUVALIIESERVALYASNPMEGLYGLNLSER 220
DB 693 CCGCTAAGATCTCTTCTATATCTGATATCTCAGTGAAGACATGGAGGCGCAGAT 752
QY 221 GLUASNPSEHESERAPTHRTHRSERVALSPILEILEVALIHRGJUNILETRPLY 240
DB 753 GAGAAATTCCTTCAGTATACCATCTGTGATATATAGTACAGAGAAATTTGGA 812
QY 241 ALAPROLYPROVALIGLUMETVALIGUANSETHRASPPOHISPROILEYSLIETHR 260
DB 813 GCACTCAAAACCTGTGAGATGTGGAACCAATCAACGATCTCAACCCCAATCAACACT 872
QY 261 GNLVALARGTRPASNAPPROGLYVALIAGINTYRSEULEUVALASPLVGLNLSLEUPRO 280
DB 873 CAGGTCCGCTGGAATGATCCCGGTGACAAATATTCCTTAGTTACAAAGAGAGCTGCCA 932
QY 281 ARGPHROPHSESERILEASPGINGLUGLYASPILEYRVALIHRGINPROLEUASPARG 300
DB 933 AGATTCCTCATTTTCAATTTGACAGAGAGGATATTTACGTGACTCAGCCCTTGACCGA 992
QY 301 GLUGLULYASPALAITYRVALIPIETYRVALAVALALYASPGIUTYRGLYLYSPROLEU 320
DB 993 GAAGAAAAGATGCAATGTTTTTATGCAGTTCGAAAGATGAGTACGGAACCAACACTT 1052
QY 321 SETYRPROLEUGLUILEHISVALYASPILEYASPILEASNPASNPPOPHRTHRYS 340
DB 1053 TCAATATCCCGTGAATTCATGTAAGTTAAAGATTAATATATATCACTACATG 1112
QY 341 PROSERPROVALIHRVALPHEGLUVALIGINGUANGLUARGLEUGLYANSERIEGLY 360
DB 1113 CCGTACCAAGTAACCTATTTGAGTCCAGAGAAATGAACGATGGGTAAAGTATCCG 1172
QY 361 THRLEUTHRALAHISAPARGAPGLUGLUBSERTHRALASNPSEHLEUANTYRARG 380
DB 1173 ACCCTTAACGACATACAGAGGATGAAGAAATATCCGCAACGTTTTCTAAACTACAG 1232
QY 381 ILEVALIGLUGLINTHRPROLYSLEUPROMETASPGIYLEUPHEULIIEGLNTHRYRALA 400
DB 1233 ATTTGTGAGAACAACTCCAAACTCTCCATGATGACCTCTTCAATCAACAACTATGCT 1292
QY 401 GLYMETLEUGNLEUALYSGNLSERLEYSLYGLNAPTHRPROGLINTYRASNLEU 420

DB 1293 GGAATGTACAGTTAAGTAAACAGTCTTGAGAGAACAGATACTCTCAGTACACTTA 1352
QY 421 THRILEGLUVALSERSPYASAPHELYSTHLEUCYEPHEVALIGNILEASNPVALL 440
DB 1353 ACCATAGAGGTGTCTACAAAGATTTTCAGAACCTTTGTTTGTGCAAAATCAACGTTAT 1412
QY 441 ASPILASNPASNPINILEPROILEPHEGLULYSESERASPTRGLYANLEUTHRLEUALA 460
DB 1413 GATATCAATGATAGACCCCATCTTTGAAAAATCGATTATGAAACCTGACTCTTGCT 1472
QY 461 GLUASPTHRASNILEGLYSETHRIIELEUTHRIEGLNALATHRASPALASPGIUPRO 480
DB 1473 GAAGACAAACACTTGGGTCCACCATCTTAACATCAAGCCACCTGATGCTGATGAGCCA 1532
QY 481 PHETHRGLYSESERLEYSLILEUTHYRHSILEILEYGLYASNPSEGLUGLYARGLEU 500
DB 1533 TTTACTGGAGGATTCATAAATTCGTATCATCATTAAGGAGACGTAGAGGACCCCTG 1592
QY 501 GLYVALASPTHRASPPOHISTHRASNTHRGTYRVALIILELYSLYSPROLEUASP 520
DB 1593 GGGGTGACACAGATCCCATACCAACACCGATATGTCATTAATAAGCTCTTGAT 1652
QY 521 PHEGLUTHRALAVALSERASNILEVALIPIEYSLAIGLUASNPPOGLUPROLEUVAL 540
DB 1653 TTTGAACAGACAGCTGTTTCCAACATGTGTTCAAGGAGAAATCTGAGCCTTAAGT 1712
QY 541 PHEGLYVALIYSTRASNALASERSEHPHEUALYSPHETHRLEUILEVALIHRASPVAL 560
DB 1713 TTTGGGTAAAGTACAAATCAAGTTCTTTGGCAAGTTCAAGTTATGTCAGATGTG 1772
QY 561 ASNGVALAPROGLNPHSESERGINHISVALIPHEGLNALYSEVALSERGIUNASPVALLA 580
DB 1773 AATGAAGCACTCAATTTTCCACACGTAATCCAAAGCAAAAGTCAAGTACATGATGCT 1832
QY 581 IIEGLYTHRLYSEVALIYASNPVALIHRALYASNPPOGLUGLYLEUASPILESERY 600
DB 1833 ATAGGCACTAAAGTGGGCAATGACTGCCAAGATCCAGAAAGTCTGACATTAAGCTAT 1892
QY 601 SERLEUARGLYASPTHRRARGLYTRPLEULYSLIEASPIHISVALIHRGLIULIEPHE 620
DB 1893 TCACTGAGGGGAGACACAAAGATGGCTTAAATTTGACACAGTCACTGTGATGATCTT 1952
QY 621 SERVALALAPROLEUASPARGLUVALIAGYSESPROTYRARGVALIGNVALAIAATHR 640
DB 1953 AGTGTGCTCCATTTGACAGAGAGCGGAAGTCCATATCGGTACAGTGGGCGCACAA 2012
QY 641 GLUVALIGLYSESERSEUSERSEVALISERGIUPHEHISLEULEUMETASPVAL 660
DB 2013 GAAGTAAGGGGGTCTTCTTGAGCTGTGTGACAGTTCCACTATCTTTATGATGTG 2072
QY 661 ASNPASNPPOPHRARGLEUALALYASPILEYRTHRGLYLEUPHEPHECYHSIPROLEU 680
DB 2073 AATGACAAACCTCCAGGCTACCAAGGACTACAGGGGCTTGTCTTCCCATCCCCCTC 2132
QY 681 SERALAPROGLYSESERLEULEPHEGLUVALIHRASPAASNPINISLEUPHEARGLY 700
DB 2133 AGTGACCTGGAAGTCTCATTTTCAGAGCTAGTAGATGACACTTATTTGGGGGT 2192
QY 701 PROHISPHETHRPHSESERLEUGLYSEERGLYSELEUGNANASNPTRGILUVALISERY 720
DB 2193 CCCCATTTTACATTTTCCCTCGCAGTGAACCTTACAAAGACTGGGAATTTCCAAA 2252
QY 721 ILEASNGLYTHRHIALAARGLEUSERTHRARGHISHRASPHNEGLUVALARGALAYR 740
DB 2253 ATCAATGTGACTCAAGCCGACTGTCTACAGGCAACAGAGTTTGAGAGAGGAGATAT 2312
QY 741 VALIALLIENILEARGIILEASNPSPGLYIARGPPOPHROLEUGLUGLYILEVALISERLEU 760
DB 2313 GTGCTCTGATCCGATCAATGATGGGGTGGCCACTTGGAAAGCAATGTTCTTTA 2372
QY 761 PROVALIHRPHECYSECYSEVALIGLUGLYSECYSPHEARGPROALAGLYHISGLINTHR 780

Db 2373 CCAGTTACATTCGACGTTGTGNGAAGAAAGTTGTTCCGGCAGAGCTCACCAACT 2432
 QY 781 GYIIEPProThValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGly 800
 Db 2433 GGGATACCCCACTGGGCGATGAGGATGTTACTGCTGACCAACCCCTTCTGGGATGGT 2492
 QY 801 ILEIIELeuAlaValAlaPheIleArgIleLeuLeuAspIleGlyLeuAspAlaValGlu 820
 Db 2493 ATATATTTTCCAGTTGGTTTATCCGATTAAGAGATTAAGCAAGATTAATTTGAA 2552
 QY 821 SerAlaGlnAlaSerGluValIleProLeuArgSer 832
 Db 2553 AGTGCTCAAGCATCTGAAGTCAACCTCTGAAGC 2588
 RESULT 8
 ABL67667
 ID ABL67667 standard; DNA: 3654 BP.
 AC ABL67667;
 DT 15-MAY-2002 (first entry)
 XX
 DE Oesophagus cancer related gene sequence SEQ ID NO:6004.
 KW Human, cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW Cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 OS Homo sapiens.
 XX
 PN M0200194629-A2.
 XX
 PD 13-DEC-2001.
 PF 30-MAY-2001; 2001MO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-231133P.
 PR 18-SEP-2000; 2000US-231617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 20-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.
 PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 28-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR

PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAIL-) AVALON PHARM.
 XX
 XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 DR WPI; 2002-188264/24.
 XX
 PT Screening for anti-neoplastic agent involves exposing cells to a
 PT chemical agent to be tested for anti-neoplastic activity, and
 PT determining a change in expression of a gene of a signature gene set -
 XX
 PS Claim 1; SEQ ID 6004; 44pp; English.
 XX
 CC The present invention describes a method (M1) for screening for an
 CC anti-neoplastic agent. The method involves exposing cells to a chemical
 CC agent to be tested for anti-neoplastic activity, determining a change in
 CC expression of at least one gene (I) of a signature gene set, where (I)
 CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 CC to ABL70110), or is at least 95% identical to (S), where a change in
 CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
 CC activity and can be used in gene therapy. M1 can be used for screening
 CC an anti-neoplastic agent, and can be used for producing a product which
 CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilm's tumour.
 XX
 SQ Sequence 3654 BP; 1052 A; 804 C; 777 G; 1021 T; 0 other;
 Alignment Scores:
 Pred. No.: 0 Length: 3654
 Score: 4303.00 Matches: 828
 Percent Similarity: 99.76% Conservative: 2
 Best Local Similarity: 99.52% Mismatches: 0
 Query Match: 24 Indels: 0
 DB: Gaps: 0
 US-10-025-380-1081 (1-832) x ABL67667 (1-3654)
 QY 1 MetIleLeuGlnAlaHisIleuHisSerLeuCyLeuLeuMetLeuTyrLeuAlaThrGly 20
 Db 93 ATATATCTTCAAGCCATCTTCACTCCCTGTGTCTTATATGTTATTTGGCACTGGA 152
 QY 21 TyrGlyGlnGluPheSerGlyPheSerGlyProLeuLeuProMetThrPheSerIleTyrGlu 40
 Db 153 TATGCCCAAGAGGGAGAGTTAGTGAACCCCTGAACCATGACATTTCTATTATGAA 212
 QY 41 GlyGlnGluPheSerGlnIleIlePheGlnPheLeuAlaAspProAlaValIlePhe 60
 Db 213 GGCCCAAGAACCGAGCAATTAATATTCAGTTTAAGGCCAATCTCTGCTGACCTTT 272
 QY 61 GlnLeuThrGlyGlnThrAspAlaIlePheValIleGlnArgGlnGlyLeuLeuTyrTyr 80
 Db 273 GAACCTAACCTGGGAGACAGCAACATATTTGTATGAAGGAGGACCTTCTGTATTAC 332
 QY 81 AsnArgAlaLeuAspArgGlnThrArgSerThrHisAsnLeuGlnValAlaAlaLeuAsp 100
 Db 333 AACAGACCTTGGACAGGAGAAACAATCTACTACATCTCCAGTTGACGCCCTGGAC 392
 QY 101 AlaAsnGlyIleIleValaGlnGlyProValProIleThrIleGlnValIleAspIleAsn 120
 Db 393 GCTAATGGAATTAATAGGAGGGGTCCAGTCCCTATCAACCATTAAGTGAAGACATCAAC 452

QY 121 AspaAnaIrgProThrPheLeuGlnSerIysTYrGlnGlySerValaIArgGlnAsnSerArg 140
Db 453 GACAAATGACCCAGCTTCTCCAGTCAAGTAAGAGGCTCAGTAAGGCAAACTCTCGC 512
QY 141 ProGlyIysProPheLeuTYrValaAsnAlaThraAspLeuAspProAlaThrProAsn 160
Db 513 CCAGAAAGCCCTCTTGATGTAATGCCACAGACTGATGATCCGGCCACCTCCAAAT 572
QY 161 GlyGlnLeuTYrTYrGlnIleValaIleGlnLeuProMetIleAsnAsnValMetTYrPhe 180
Db 573 GGCCAGCTTTATTCAGATGTTGATCCAGCTTCCATGATCAAAATGTCATGTAATT 632
QY 181 GlnIleAsnAsnIystrGlyValaIleSerLeuThraIrgGlnGlySerGlnGlnLeuAsn 200
Db 633 CAGATCAACAACAAAGCGGAGCATCTCTTACCCGAGAGGATCTCAGGAATGAT 692
QY 201 ProAlaIysAsnProSerTYrAsnLeuValaIleSerValaIysAspMetGlyGlnSer 220
Db 693 CCGCTTAAGATCCTTCCATTAATCTGGTATCTCAGTGAAGGACATGGGAGCCAGAGT 752
QY 221 GluAsnSerPheSerAspThrThraSerValaAspIleIleValaThrgluAsnIleTrpIys 240
Db 753 GAGATTCCTTCAGATGATACCACTCTGGATATCATGACAGAGAAATATTGGA 812
QY 241 AlaProIysProValaGlnMetValaGlnAsnSerThraAspProHisProIleIysIleThr 260
Db 813 GCACCAAAACCTGGAGATGGATGGAAAACCACTATCTCTCACTCCCATCAAAATCACT 872
QY 261 GlnValaIrgTrpAsnAspProGlyValaGlnIySerLeuValaAspIyGlnIysLeuPro 280
Db 873 CAGGTGGGTGGATGATCCCGGTGCACAATATCTTGAAGTGAAGAGAGAGCTCCA 932
QY 281 ArgPheProPheSerIleAspGlnGlnGlyAspIleTYrValaThrglnProLeuAspArg 300
Db 933 AGATTCCTCATTTCAATTGACCAAGAGAGATATTAAGTACTGACCCCTTGAGACGA 992
QY 301 GlnGlnIyIysAspAlaTYrValaPheTYrAlaValaIalysAspGlnIyTYrGlyIysProLeu 320
Db 993 GAAGAAAGATGATATGTTTTTATGCAAGTTCGCAAGGATGATGAGGAAACCACTT 1052
QY 321 SerTYrProLeuGlnIleHisValaIysValaIysAspIleAsnAspAsnProProThrCys 340
Db 1053 TCATATCCGCTGGAATTCATGTAAAGTTAAAGATTTAATGATATCCACCTACATGT 1112
QY 341 ProSerProValaThraValaPheGlnValaGlnGlnAsnGlnArgLeuGlnIysAsnSerIleGly 360
Db 1113 CCGTCACAGATACCGATTTGAGGTCCAGAGATGAACGACTGGGTAAACAGATTCGGG 1172
QY 361 ThraLeuThraIalHisAspArgAspGlnGlnAsnThraIalAsnSerPheLeuAsnTYrArg 380
Db 1173 ACCCTTACTCAGATGACAGGATGAGAGAAATATCTCCAAAGCTTTCTTAACTACAGG 1232
QY 381 IleValaGlnIleThraProIysLeuProMetAspGlyLeuPheLeuIleGlnIyTYrAla 400
Db 1233 ATGTGGAGCAAACTCCCAAACTTCCATGATGATGATCTTCTTAACTCAAACTATGCT 1292
QY 401 GlyMetLeuGlnLeuAlaIysGlnSerLeuIyIysGlnAspThraProGlnIyTYrAsnLeu 420
Db 1293 GGAATGTTACAGTTAGTAAACAGTCTTGAAGAGAAAGATATCTCTCGTCAAACTTA 1352
QY 421 ThrIleGlnValSerAspIysAspPheIysThraLeuCysPheValaGlnIleAsnValaIle 440
Db 1353 ACGATAGAGGTGCTGCAAAAGATTTCAAGACCTTTGTTTGTGCAAACTCAAGCTTAT 1412
QY 441 AspIleAsnAspGlnIleProIlePheGlnIyIysSerAspTYrGlyAsnLeuThraLeuAla 460
Db 1413 GATATCATATGATGACCCCATCTTTGAAATAATCAATATGGAACCTGACCTTGTCT 1472
QY 461 GluAspThraAsnIleGlySerThraIleLeuThraIleGlnAlaThraAspAlaAspGlnPro 480
Db 1473 GAAGACACAAACATTGGGTCCACCATCTTAAACATCCAGGCCACTGATGATGAGCCA 1532

QY 481 PheThrGlySerSerIysIleLeuTYrHisIleIleIysGlyAspSerGlnGlyArgLeu 500
Db 1533 TTACTGGGAGATTCTAAATTTCTGTATCATATCATAAAGGAGACAGTGAAGGACGCTG 1592
QY 501 GlyValaAspThraAspProHisThraGlnThrglyTYrValaIleIleIysIysProLeuAsp 520
Db 1593 GGGGTGACACAATATCCCTTACCAACACCGGATATGATCAATTAATTAAGGCTCTTGAT 1652
QY 521 PheGlnThraIalAlaIysAsnIleValaPheIysAlaGlnAsnProGlnProLeuVala 540
Db 1653 TTTCGAACAGCAGCTGTTTCCATCATTTGCTTCAAGGCAAAATCTGAGCTTATG 1712
QY 541 PheGlyValaIysTYrAsnAlaSerSerPheAlaIysPheThraLeuIleValaThraAspAla 560
Db 1713 TTGTGTGAAGTACATGATGCAATGTTTGGCCAACTGTCAGCTTATTTGACAGATGTG 1772
QY 561 AsnGlnAlaProGlnPheSerGlnHisValaPheGlnAlaIysValaSerGluAspValaAla 580
Db 1773 AATGAAGCACCTCAATTTTCCCAACAGTATTCGAAGCGAAAGTCACTGAGATGTAGCT 1832
QY 581 IleGlyThraIysValaGlyAsnValaThraAlaIysAspProGlnIyLeuAspIleSerTYr 600
Db 1833 ATAGGACTTAAGTGGCAATGTGACTGCCAGAGATCCAGAGGCTTGACATPAAGCTAT 1892
QY 601 SerLeuArgGlyAspThraIrgGlyTYrPLeuIyIleAspHisValaThrglyGlnIlePhe 620
Db 1893 TCATGAGGGGAGACACAAGAGGTTGGCTTAATAATGACACGCTGATGATCTTT 1952
QY 621 SerValaIalProLeuAspArgGluAlaGlySerProTYrArgValaGlnValaIalThr 640
Db 1953 AGTGTGGCTTCATTTGACAGAGAAAGCGGAAGTCCATATCGGTACAGAGTGGGCGACA 2012
QY 641 GluValaGlyIysSerSerLeuSerSerValaSerGlnPheHisLeuIleLeuMetAspVala 660
Db 2013 GAGCTAGGGGGGCTTCTTCTGAGCTGTGTCTGAGTCTCCACTGATCTTATGAGATGTG 2072
QY 661 AsnAspAsnProProArgLeuAlaIysAspTYrThrglyLeuPhePheCysHisProLeu 680
Db 2073 AATGACAAACCTCCAGGGGTAGCCAGCAAGACTACAGGGGTGTCTTCTGCAATCCCTC 2132
QY 681 SerAlaProGlySerLeuIlePheGlnAlaThraAspAspGlnHisLeuPheArgGly 700
Db 2133 AGTGACCTGGAGATCTCATTTTCGAGGCTACTGAGATGATCAGACTTAATTCGGGGT 2192
QY 701 ProHisPhePhePheSerLeuGlySerGlySerLeuGlnAsnAspTrpGlnValaSerIys 720
Db 2193 CCCCATTTTACATTTTCCCTCGGCACTGGAAGCTTCAAAACGACTGGGAAGTTTCCAA 2252
QY 721 IleAsnGlyThraHisAlaIrgLeuSerThraArgHisThraAspPheGlnGlnArgAlaTYr 740
Db 2253 ATCAATGTATCTATATCCGACTGTCTTACCAAGCAACAGATTTTAAGAGAGGAGTAT 2312
QY 741 ValValleuIleArgIleAsnAspGlyIyArgProProLeuGlnGlyIleValaSerLeu 760
Db 2313 GTGCTTGTATCCGATCAATGATGGGGGTGGCCACCTTGGAAAGCATTTGTTCTTA 2372
QY 761 ProValaThraPheCysSerCysValaGlnGlySerCysPheArgProAlaGlyIysGlnIle 780
Db 2373 CCAATTACATCTGCAAGTGTGTGGAAGAAAGTTGTTCCGGGCAAGGCTCCAGACT 2432
QY 781 GlyIleProThraValaGlyMetAlaValaGlyIleLeuLeuThraThraLeuValaIleGly 800
Db 2433 GGGATACCACTGTGGGCAATGGAGATGGATATCTGTCAACCACTTCTGTGATTTGT 2492
QY 801 IleIleLeuAlaValaValaPheIleArgIleIysIysAspIysGlyIysAspAsnValaGlu 820
Db 2493 ATATATTTAGCAGTTGTGTTATCCCATTAAGAAAGATPAAGAGGCAAAATATATGTTGA 2552
QY 821 SerAlaGlnAlaSerGlnValaIysProLeuArgSer 832
Db 2553 AGTGTCAAGCATCTGAAGTCAAACTTGAAGAGC 2588

RESULT 9

AAH98190
 ID AAH98190 standard; cDNA; 2820 BP.
 AC AAH98190;
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST-derived coding sequence SEQ ID NO: 47.
 XX
 KM Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KM diagnostics; forensic test; gene mapping; genetic disorder;
 KM biodiversity; gene therapy; nutrition; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V,
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 DR WPI; 2001-476164/51.
 DR P-PSDB; AAM23531.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -
 XX
 PS Claim 1; Page 213-214; 1275bp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention.
 XX
 SQ Sequence 2820 BP; 593 A; 867 C; 821 G; 539 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1.19e-70 Length: 2820
 Score: 922.50 Matches: 237
 Percent Similarity: 51.02% Conservative: 163
 Best Local Similarity: 30.23% Mismatches: 353
 Query Match: 21.35% Indels: 31
 DB: 22 Gaps: 19
 US-10-025-380-1081 (1-832) x AAH98190 (1-2820)
 QY 62 LeuHrGlyGlu-----ThrAspAnIlePheValIleGluArgGlu---Gly 76
 DB 220 CTGTCAGGGGACTCAGGCAAGGCACTGAGGCCCATTTGATGCCAGATTCTGGC 279
 QY 77 LeuLeuTYrTYrAsnATgAlaLeuAspArgGluThrArgSerThrHISAsnLeuGlnVal 96
 DB 280 TTCCTGCTGTCGACAGGGCCCTGAGCCGAGAGAGAGGAGTACACACTACAGGTC 339
 QY 97 AlAlaLeuAspAlaAsnGlyTlleValGluGlyProValProIleThrIleGluVal 116
 DB 340 ACCCTGAGATGACAGATGACATGTCTTGCGGGGTCCACAGCCTGTGCTGCACGTG 399

QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerLYrGlySerValArg 136
 DB 400 AAGGATGAGATGATACAGAGTGGCCCATTTCTCTCAAGCATCTACAGAGTGGCTGAGC 459
 QY 137 GlnAsnSerArgProGlyLYrProPheLeuTYrValAsnAlaThrAspLeuAspPro 156
 DB 460 CGGGTACCAAGGCTGGCATCCCTTCCTTCCTTGGAGCTTCACAGGGGATGAGCCA 519
 QY 157 AlaThrProAsnGlyLysLeuTYrTYrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 520 GGCACAGCCACTCCGATCTTCATTCACATCTCAAGCTCAGGCTCAGCCGCTTCC 579
 QY 177 ValMetTYrPheGlnIleAsnAsnLYrThrGlyAlaIleSerLeuThrArgGlySer 196
 DB 580 CCAGACATGTTCCAGCTGAGCTGGAGCTGGGGGCTCTGGCCCTCAGCCCAAGGGAGC 639
 QY 197 GlnGluLeuAsnProAlaLYrAsnProSerTYrAsnLeuValIleSerValLYrAspMet 216
 DB 640 ACCAGCCTTGACCAACCCCTGAGAGAGACTTACAGCTGTGTACAGGTCAAGACATG 699
 QY 217 GlyGlyLysSerGluAsnSerPheSerAspThrThrservalAspIleValThrGln 236
 DB 700 GTGACCAAGGCTC--TCAGGCCACAGGCCACTGCCCAGTGAAGTCTTCATCATAGAG 756
 QY 237 AsnIleTrpLYrAlaProLYrProValGluMetValGluAsnSerThrAspProHISPro 256
 DB 757 AGCAGCTGGGTGTCCTCAGAGCTTATCCAGCTGGAGAGATCTCAAGTCTATACCCG 816
 QY 257 IleLYrIleThrGlnValArgTrpAsnAspProGlyAlaGlnTYrSerLeuValAspLYs 276
 DB 817 CACCACATGGCCAGGAGTACCTGAGTGGGGTGAATGTGCATCATACCTG----- 867
 QY 277 GlnLYrLeuProArgPheProPheSerIleAspGlnGluGlyAspIleTYrValThrGln 296
 DB 868 GAGAGCCATCCCGGGGACCTTTGAAGTAATCAGAGGAACTCTTACGACACAGA 927
 QY 297 ProLeuAspArgGluGluLYrAspAlaTYrValPheTYrAlaValAlaLYrAspGluTYr 316
 DB 928 GACTGGACGAGAGAAGCGGCTGAGTACCTGCTCCAGGTGGGGCTCAGATTCCCAT 987
 QY 317 GlyLYrProLeuSerTYrProLeuGlnIleHISValLYrValLYrAspIleAsnAspAsn 336
 DB 988 GGGAGAGACTATCCGGCCCTCTGAGCTGCAGCTGTGTATGATGAGATGACACAC 1047
 QY 337 ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly 356
 DB 1048 GTGCCTATCTGCTCCCTCCCGCTGACCCCAAGTACGATCTTACGCTCAGTCCACAGGT 1107
 QY 357 AsnSerIleGlyTYrLeuThrAlaHISAspArgAspGluGluAsnThrAlaAsnSerPhe 376
 DB 1108 ACTGAGTACTGACTGCTGCTGACAGAGATGAGATGCCCGCGGCTCCCAATTCCAC 1167
 QY 377 LeuAsnTYrArgIleValGluGlnThrProLYrLeuProMetAspGly---IleuPheLeu 395
 DB 1168 GTTGTGATACGCTCTGAGCCCTGAGCCCTGAGAGATGAGGTAGAGGAGAGGCTTCAG 1227
 QY 396 IleGlnIleTYrAlaGlyMetLeuGlnLeuAlaLYrGlnSerLeuLYrLYs---GlnAsp 414
 DB 1228 GTGAGCCCACTTCAGAGCTGAGCTGGGGGTGCTCCCATCTCCAGAGAGCCAGAAC 1287
 QY 415 ThrProGlnTYrAsnLeuThrIleGluValSerAspLYrAsp-----PheLYrThrLeu 432
 DB 1288 ATCTGCTTGTGCTGCTGGCCATGAGACTGGCAGAGGGGAGGGGCTTCACAGACAGC 1347
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLYrSer 452
 DB 1348 TGTGAAGTCAAGATCCAGTCAAGATATCAATGATCAACGCCCTGATGTTCACTTCC 1407
 QY 453 AspTYrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1408 CAGATTGGGCTTATAGCTCTCCCTGAGAGATGAGAGCCGGGAGCTGTGGGCACTGA 1467
 QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLYrIleLeu---TYrHIS 490


```

Db      1468 AACGCCCATGATGCTGACCTGAGCC-----GCCTTCGCCCTCATGATTTTGCC 1518
Qy      491 11ellelveglYAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
Db      1519 ATTGAGAGGGGAGACACAGAGGAGCTTTTGCTGGATTGGAGCCA-----GACTCT 1572
Qy      511 G1YTYrYVal11ellelYs-----LysProLeuAspPheGluThrAlaAlaValSerAsn 528
Db      1573 GGGCATGTTAGACTCAGACTCTGCAAGAACTGATTAGAGGAGCTCCAAAGTCATGAG 1632
Qy      529 11eValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTYrAsnAlaSer 548
Db      1633 GTGTGTGTGTGTGTGACAGAGTGTGCGCAAGCTGTGTG---GGGCCAGGCCCAAGCTTGA 1689
Qy      549 SerPheAlaLysPheThrLeu11eValThrAspAlaAsnGluAlaProGluPheSerGln 568
Db      1690 GCCACGCCCAAGCTGATCTGTGCTAGTGAGAGAGTGAATGCCACCCCAAGTTGGACCA 1749
Qy      569 HisValPheGluAlaLysValSerGluAspValAla11eGlyThrLysValGlyAsnVal 588
Db      1750 GAGAGCTACAGAGGCGCATGTGCTCCATCATGTGCCCCAGCCGCTCTTCCTGTGATCATC 1809
Qy      589 ThrAlaLysAspProGluGlyLeuAspL11eSerTYrSerLeuArgGlyAspThrArgGly 608
Db      1810 CAGGCTCCCAAGCCCATCAAGCCCAAGCCCTCAGTTCTCCCTAGTCAATGACTCAGAGGCG 1869
Qy      609 ThrLeuLys11eAspHisVal11eThrGlyL11ePheSerValAlaProLeuAsp---Arg 627
Db      1870 TGGCTCTGCAATTGAAATTTCTCCGGAGAGTGACACCGCCAGCTCCCTGACAGGCGGCC 1929
Qy      628 GluAlaGlySerProTYrArgValGlyValAlaThrGluValGlyGlySerSerLeu 647
Db      1930 CAGCTTGCGGAGACCTACAGCTGCTTGTGGAGGCCAGATACAGATGAGCCGAGACTG 1989
Qy      648 SerSerValSerGluPheHis11eLeu11eLeuMetCaspValAsnAspAsnProProArgLeu 667
Db      1990 AGGCTTCTGACACCCCTGTGATCTCACTTAAAGGCCCTCTCCCTGACAGCCCTGACT 2049
Qy      668 AlaLysAspTYrThrGlyLeuPhePheCysHisProLeuSerAlaProGlySerLeu11e 687
Db      2050 CTTCGCCCTTGCTCCCAATACCTCTGTGACACCCGCCAAGACCATGCTGTGATGTG 2109
Qy      688 Phe---GluAlaThrAspAspAspGlnHis11eLeuPheArgGlyProHisPheThrPheSer 706
Db      2110 AGTGAACCCAGCAAGACCCCGAATCTGGCCAGTGGGACGCTCC---TACAGCTTACC 2166
Qy      707 LeuGly---SerGlySerLeuGluAsnAspTYrGluValSerLys11eAsnGlyThrHis 725
Db      2167 CTTCGTCCCAACCCCAAGCTGCAACGGGATTGGCGCTCCAGACTCTCAATGGTTCCAT 2226
Qy      726 AlaArgLeuSerThrArgHis11eThrAspPheGluGluArgAlaTYrValValLeu11eArg 745
Db      2227 GCTTACTCTACCTTGCCCTTGCTGATGGGTGAGCAGCTAACAACATAATCCCGTGTG 2286
Qy      746 11eAsnAspLysGlyArgProProLeuGluGly11eValSerLeuProVal11ePheCys 765
Db      2287 GTGAGCAACAATGCCAG-----ATGTGGAGAGCTCTGTTCAGAGTCAATGCTGTCT 2337
Qy      766 SerCys---ValGluGlySerCysPheArgProAlaGlyHisGlnThrGly11eProThr 784
Db      2338 CGCTGCAACCTGAGGGGAGGAGTGCATGCGCAAGGTGGCGCATGAAGGGCATGCCACG 2397
Qy      785 ValGlyMetValValGlyLysLeuLeuThrThrLeuLeuVal11eGlyLys11eLeuAla 804
Db      2398 AAGCTGTGGAGGGGAGCTCTTGTAGGCAACCTGTGATCAATAGGAATCTTCTCATC 2457
Qy      805 ValValPhe11eArg11eLysLysAspLysGlyLysAspAsnValGluSerAlaGlnAla 824
Db      2458 CTGATTTTACACCACTGACCATGTCAAGAAAGAACCCCGATCAACAGACAGACAGC 2517
Qy      825 SerGluValLys 828

```

```

Db      2518 GTGCCCTGGAAG 2529
RESULT 10
AAH57507
ID      AAH57507 standard; cDNA; 2820 BP.
XX
AC      AAH57507;
XX
DT      10-SEP-2001 (first entry)
XX
DE      Human kidney cell specific cDNA sequence SEQ ID NO:347.
XX
KW      Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KM      lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KN      metabolic disease; developmental disease; cytostatic; immunomodulatory;
KV      neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS      Homo sapiens.
XX
PN      WO200132927-A2.
XX
PD      10-MAY-2001.
XX
PF      02-NOV-2000; 2000MO-US30396.
XX
PR      04-NOV-1999; 99US-0163508.
XX
PA      (INCY-) INCYTE GENOMICS INC.
XX
PI      Sornasse T, Seilhammer JU, Watson GA;
XX
DR      WPI; 2001-291057/30.
XX
PT      New cell and tissue specific polynucleotides useful for diagnosis,
PT      prognosis or monitoring of treatments for disorders where the gene is
PT      associated with a cancer, immunopathology or neuropathology -
XX
XX      Claim 1; Page 264; 327pp; English.
XX
CC      AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC      sequences (I). (I) can have cytosolic, immunomodulatory and
CC      neuroprotective activities, and can be used in gene therapy. (I) and
CC      proteins (II) encoded by them are used in high throughput screening
CC      assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC      mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC      their fragments, immunoglobulins, inhibitors, drug compounds and
CC      pharmaceutical agents. Expression of (I) in a sample indicates the
CC      differentiation of embryonic stem cells into a tissue selected from
CC      brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC      tissues. (I) and (II) are used to produce an expression profile that
CC      defines a metabolic or developmental process, treatment, condition,
CC      disease or disorder. The gene profile can be used for diagnosis,
CC      prognosis or monitoring of treatments and for investigating a
CC      predisposition to a disorder where the gene is associated with a
CC      cancer, immunopathology or neuropathology.
XX
SO      Sequence 2820 BP; 586 A; 869 C; 824 G; 541 T; 0 other;

Alignment Scores:
Pred. No.:      1.19e-70      Length:      2820
Score:          922.50      Matches:      237
Percent Similarity: 51.02%      Conservative: 163
Best local Similarity: 30.23%      Mismatches: 353
Query Match:      21.35%      Indels:      31
DB:              22      Gaps:      19

US-10-025-380-1081 (1-832) x AAH57507 (1-2820)
Qy      62 LeuThrGlyGlu-----ThrAspAsn11ePheVal11eGluArgGlu---Gly 76
Db      221 CTGTACGGGAGACTCAGCAAGCAACTGAGGAGGCCCAATTTGCTATGATGATCAAGATTCCTGCG 280
Qy      77 LeuLeuTYrTYrAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGluVal 96

```

```
Db 281 TTCTGCTGGACGACGAGCCCTGGACCGAGAGAGACGAGCAGTACAGCTACAGTCC 340
Qy 97 ALAAlaLeuAapAlaAnglyLeIleValGluGlyProValProIleThrIleGluVal 116
Db 341 ACCGTGAGATGACGATGATGATCTTGTGGGGTCCACAGCTGTGCTGTGACAGTGTG 400
Qy 117 LysAapIleAapAapAapAapProThrPheLeuIleSerIleValGluGlySerValArg 136
Db 401 AAGATAGAAATGACAGGTGCTTCATCTTCACAGCATCTACAGAGCTGGCTAGC 460
Qy 137 GluAapSerArgProGlyIleValProPheLeuValAlaAapIleAapLeuAapAapPro 156
Db 461 CGGGGTACAGAGCTGGGATCCCTTCCTCTCTGAGGCTTACAGACGGGATGAGACA 520
Qy 157 AlaThrProAanglyGluLeuValIleValIleGluLeuProMetIleAapAap 176
Db 521 GGCACACCAACTGGATCTTCATCCACATCTGAGCCAGGCTCAGCCAGCCCTTCC 580
Qy 177 ValMetIlePheGluIleAapAapIleThrGluValIleSerLeuThrArgGluGlySer 196
Db 581 CCAGACATGTTCACGCTGAGGCTCGGCTGGGGCTCTGAGCCCTGAGCCCAAGGGAGC 640
Qy 197 GluIleLeuAapProAlaValAapProSerIleAapLeuValIleSerValIleAapMet 216
Db 641 ACCAGCCTTGACACGCTCGGAGAGACCTACAGCTGTGATGACAGTCAAGATGAGATG 700
Qy 217 GlyIleGluIleAapSerPheSerAapThrThrSerValAapIleIleValIleGlu 236
Db 701 GGTACACAGGCC--TCAGGCCACACAGGCCACCTGCCAGGAGGCTTCATCATGAG 757
Qy 237 AapIleThrIleAapProValProValGluMetValGluAapSerIleAapProIleAap 256
Db 758 AGCAGCTGGGTGCTCTAGAGCTTACCACTGGGAGAGAACTCAAAAGTCTTATCCG 817
Qy 257 IleValIleThrGluValArgIleAapAapProGlyAlaGluIleValIleValAap 276
Db 818 CACACACATGGCCACGATCACTAGAGGGGGTATGACATCACTG----- 868
Qy 277 GluIleLeuProAapPheProPheSerIleAapGluGluAapIleValIleThrGlu 296
Db 869 GAGAGCAGATCCCGGGGACCTTTGAGTGAATGACAGAGGAAACCTTACAGTACAGAGA 928
Qy 297 ProIleAapArgGluGluValAapAlaThrValIleThrValAlaValAapAapGlu 316
Db 929 GAGCTGACAGAGAGCCAGGCTGAGTACCTGCTCAGAGTGGGGCTCAGAAATCCCAT 988
Qy 317 GluIleAapProLeuSerIleValGluIleValIleValIleValIleValAapAap 336
Db 989 GGGAGAGACTATGGGGCCCTCTGAGCTGACAGCTGCTGATGATGAGATGAGATGAGACA 1048
Qy 337 ProProThrIleValProSerProValIleValIleGluValIleGluAanglyGly 356
Db 1049 GTGGCTATCTGCTCCCTCCCTGAGCCSACAGTCAAGATCCCTGAGTCACTGACAGCT 1108
Qy 357 AapSerIleGlyIleThrIleThrAlaIleAapAapArgGluAapIleAapAapPhe 376
Db 1109 ACTGAGAGTACATCTCTCAGAGAGATGACAGATGAGTCCCGGGCTCCCAATTCAC 1168
Qy 377 LeuAapIleArgIleValGluIleThrProIleAapProMetAapGlyIleAapPhe 395
Db 1169 GTTGTGATACGCTCTGAGCCCTGAGCTGAGATGGGGTGAAGAGGAGAGCTTCCAG 1228
Qy 396 IleGluIleThrIleGluIleGluIleValIleValIleValIleValIleValAap 414
Db 1229 GTGAGACCCCACTCAGGAGTGTGACGCTGGGGGTCTCCATCCGAGCAGGACCAAGAC 1288
Qy 415 ThrProGluIleValIleThrIleGluValIleSerAapAap-----PheIleThr 432
Db 1289 ATCTGCTTCTGCTGCTGAGATGAGCTGAGCAGGCGCAGAGGGTGGCTTACAGACAGC 1348
Qy 433 CysPheValGluIleAapValIleAapIleAapAapGluIleProIlePheGluIleSer 452
Db 1349 TGTGAATCGAAGTCCAGTACAGATATCAATGATACAGCCCTGAGTTCATCACTCC 1408
Qy 453 AapIleGlyAapIleThrIleGluIleAapIleAapIleGlySerThrIleLeuThrIle 472
Db 1409 CAGATTGGGCTTAAAGCTCTCTGAGAGTGGAGCCCGGAGCTTGGGGCATCTTA 1468
Qy 473 GluIleAapAapAapAapAapAapAapAapAapAapAapAapAapAapAapAapAap 490
Db 1469 ACAGCATTAATCTACCTCGAGGCC-----GCCCTCCGCTCATGATTTTGGC 1519
Qy 491 IleIleLeuGlyAapSerGluIleValIleValIleValIleValIleValIleVal 510
Db 1520 ATTGAAGAGGAGAACAGAGAGGAGCTTTGGCTGAGTGAAGGAGCA-----GACTTC 1573
Qy 511 GlyIleValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 528
Db 1574 GGGCATTGATGACTGACAGCTTGCAGAAACCTAGTTATGAGGAGCTCCAGTCAATGAG 1633
Qy 529 IleValPheValIleValIleValIleValIleValIleValIleValIleValIle 548
Db 1634 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1690
Qy 549 SerPheAlaIleAapPheThrIleValIleThrAapValAapGluAapProGluPheSer 568
Db 1691 GCCACCGCCACGCTGATGCTGTGATGAGAGAGTATGACCAACCCCAAGTTGAGCAG 1750
Qy 569 HisValPheGluIleValIleValIleValIleValIleValIleValIleValIle 588
Db 1751 GAGAGCTACAGGACGAGTGTCCCATACAGTCCCAAGCCGAGCTTTCCTGCTGACCATC 1810
Qy 589 ThrAlaIleAapProGluIleValIleAapIleSerIleSerIleValIleValIle 608
Db 1811 CAGCCCTCCAGACCCCATACAGCCGAGACCCCTGAGTCTTCCCTGATGATGATGAGAG 1870
Qy 609 TyrLeuValIleAapIleValIleValIleValIleValIleValIleValIleVal 627
Db 1871 TGCTCTGATGAGAAATCTCCGGGAGGTGCACACCCGAGTCCCTGAGGAGGCGCC 1930
Qy 628 GluAlaGlySerProIleValIleValIleValIleValIleValIleValIleVal 647
Db 1931 CAGCTGGGAGACCTTACAGAGTGTGTGTGAGAGCCAGAGATGACAGTACAGGAGACTG 1990
Qy 648 SerSerValSerGluPheIleValIleValIleValIleValIleValIleValIle 667
Db 1991 AGCGCTTTCAGCCCTCGATGATCACTTCAAGGCCCTGAGCCAGCCGAGT 2050
Qy 668 AlaIleAapIleThrIleValIleValIleValIleValIleValIleValIleVal 687
Db 2051 CTGACCTGCTGCTCCCATCTGATCTGACACCCCGCAGACATGAGCTTGAATCGTG 2110
Qy 688 Phe-----GluAlaIleAapAapAapAapAapAapAapAapAapAapAapAapAap 706
Db 2111 AGTGAACCCAGAGAGCCCGATCTGGCCAGTGGGACAGCTGCC-----TACAGCTTAC 2167
Qy 707 LeuGlyIleValIleValIleValIleValIleValIleValIleValIleValIle 725
Db 2168 CTGTGTCCCAACCCAGAGTGAAGAGGATGGAGCTCCAGACTCAATGATGATGATGAT 2227
Qy 726 AlaArgIleSerThrArgIleThrAapPheGluIleValIleValIleValIleVal 745
Db 2228 GCTTACTTACCTTGGCCCTGATGAGTGGTGAAGCCAGTGAACATTAATCCCGGTGTG 2287
Qy 746 IleAapAapGlyIleValIleValIleValIleValIleValIleValIleValIle 765
Db 2288 GTACAGCAATCCAG-----ATGTGACACTCTGTGTGAGATGATGATGATGATGAT 2338
Qy 766 SerCys-----ValGluIleSerCysPheArgProAlaGlyIleGluIleThrIle 784
Db 2339 CGCTGCACTGAGAGGAGGAGTCAATGAGAGTGGGCGGATGAAGGACATGCCAGC 2398
Qy 785 ValGlyMetAlaValGlyIleLeuLeuThrIleThrIleValIleGlyIleIleVal 804
Db 2399 AAGCTGCGGAGTGGGATCTTGTAGGACACCTGTGATGATGATGATGATGATGATGATG 2458
```

QY 805 ValValPheIleArgIleYsAspLySgIlySAspAenValGIuserAlaGlnAla 824
:::|||||
Db 2459 CTCATTTCACCCCACTGCAGCATGTCAAGAAAGAGCCCGATCTCAACGACGACAGC 2518
QY 825 SerGIuValYs 828
:::|||||
Db 2519 GTGCCCTGAAG 2530

RESULT 11
AAA37087
ID AAA37087 standard; cDNA; 2848 BP.
XX AAA37087;
AC
XX
XX 08-AUG-2000 (first entry)
XX
DE Human PRO1340 (UN0695) cDNA sequence SEQ ID NO:228.
XX
XX Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KM transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
ss.
XX
XX Homo sapiens.
PN M0200012708-A2.
PD 09-MAR-2000.
XX
XX 01-SEP-1999; 99WC-US20111.
PF
XX
XX 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099536.
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 15-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.

PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101745.
PR 24-SEP-1998; 98US-0101916.
PR 24-SEP-1998; 98US-0102207.
PR 24-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102310.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105593.
PR 26-OCT-1998; 98US-0105594.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0106500.
PR 30-OCT-1998; 98US-0106544.
PR 03-NOV-1998; 98US-0106556.
PR 03-NOV-1998; 98US-0106502.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 11-NOV-1998; 98US-0108775.
PR 11-NOV-1998; 98US-0108779.
PR 11-NOV-1998; 98US-0108787.
PR 11-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.

PR 18-NOV-1998; 98US-0108848.
 PR 18-NOV-1998; 98US-0108849.
 PR 18-NOV-1998; 98US-0108850.
 PR 18-NOV-1998; 98US-0108851.
 PR 18-NOV-1998; 98US-0108852.
 PR 18-NOV-1998; 98US-0108858.
 PR 18-NOV-1998; 98US-0108904.
 (GENETECH) GENENTECH INC.
 PI Baker K, Goddard A, Gurney AL, Smith V, Matanabe CK, Wood WI;
 XX WPI: 2000-237871/20.
 DR P-PSDB; AAY99405.
 XX New mammalian DNA sequences encoding transmembrane, receptor or
 PT secreted PRO polypeptides, useful for screening of potential peptide or
 PT small molecule inhibitors of the relevant receptor/ligand interactions
 XX
 PS Claim 2; Fig 131; 773pp; English.
 XX
 CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
 CC receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
 CC transmembrane and receptor PRO proteins can be used for screening of
 CC potential peptide or small molecule inhibitors of the relevant
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences
 CC encoding them have various industrial applications, including uses as
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
 CC PCR primers and hybridization probes used in the isolation of the PRO
 CC polypeptides from the present invention.
 XX
 SQ Sequence 2848 BP; 607 A; 873 C; 828 G; 540 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2.03e-69 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 DB: Gaps: 21
 US-10-025-380-1081 (1-832) x AAA37087 (1-2848)
 QY 62 LeuThrglyglu-----ThraspasnlllephValllleGluArgGlu---Gly 76
 DB 296 CTGTCAGGGAGCTCAGGCAAGCACTGAGGCCCATTTGCTATGATCCAGATTCTGGC 355
 QY 77 LeuLeuTyrrAAsnArGAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
 DB 356 TTCTCTGCTGCGAGCAGGGCCCTGGACCGAGAGACAGAGTACCACTACAGTTC 415
 QY 97 AlaAlaLeuAspAlaAsnGlylleValleGluGlyProValProIleThrIleGluVal 116
 DB 416 ACCCTGGAGATGACGATGAGCATGCTTGTGGGGTCCACAGCCTGTGCTGCACCTG 475
 QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerIleTyrrGluGlySerValArg 136
 DB 476 AAGATGATGATGACCAAGGTCCTCATTTCTCTCAAGCATCTACAGAGCTCGGTGAC 535
 QY 137 GlnAsnSerArgProGlyLysProPheLeuTyrrValAsnAlaThrAspLeuAspAspPro 156
 DB 536 CGGGTACAGGCTGCGATCCCTCTCTCTCTCTGAGGCTTCAGACGGGAGTACAGCA 595
 QY 157 AlaThrProAsnGlyGlnLeuTyrrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 596 GGCACAGCACTCGATCTTCATCCATCTCCAGCCAGGCTCCAGCCAGCTTCC 655
 QY 177 ValMetTyrrPheGlnIleAsnAsnLysThrArgAlaIleSerLeuThrArgGluGlySer 196
 DB 656 CCAGACATGTTCCAGCTGAGCCTGCGTGGGGCTTCGCGCTCAGCCCAAGGGAGGC 715
 QY 197 GlnGluLeuAsnProAlaLysAsnProSerTyrrAsnLeuValIleSerValLysAspMet 216

DB 716 ACCAGCCTTGACCAAGCCCTGAGAGACCTTACCAAGCTGTGTGATCAGTCAAGACATG 775
 QY 217 GLyGlyLysSerGluAsnSerPheSerArgThrThrSerValAspIleValThrGlu 236
 DB 776 GGTGACCAAGCC--TCAGGCCACCGAGCCATGCGACCGTGAAGTCTTCATCATAGAG 832
 QY 237 AsnIleTrpLysAlaProLysProValGluMetValGluAsnSerThrasProHisPro 256
 DB 833 AGCAGCTGGGTGTCCTTAGAGCTTACCACTGAGAGATCTCAAGGCTATACCCG 892
 QY 257 IleLysIleThrGlnValArgTrpAsnAspProGlyValAlaGlnTyrrSerLeuValAspLys 276
 DB 893 CACCAATGCGCCAGGTACAGTGGAGTGGGGTATGTGCATATACCTG----- 943
 QY 277 GlnLysLeuProArgPheProPheSerIleAspGlnGluGlyAspIleTyrrValThrGln 296
 DB 944 GAGAGCCATCCCGGAGCCCTTGAAGTAAATGCAAGAGGAAACCTTACGTGACAGCA 1003
 QY 297 ProLeuAspArgGluGlnLysAspAlaTyrrValPheTyrrAlaValAlaLysAspGluTyrr 316
 DB 1004 GAGCTGACAGAGAAAGCCAGGCTGAGTACTCTCCAGGTGGGGCTCAGAAATCCCAT 1063
 QY 317 GlyLysProLeuSerTyrrProLeuGluIleHisValLysValLysAspIleAsnAspAsn 336
 DB 1064 GCGAGAGACTATACGGCCCTCTGAGAGCTGCACGTGCTGTATGATGAGAAATGACACAC 1123
 QY 337 ProProThrCysProSerProValThrValPheGluValGlnGluAsnGluArgLeuGly 356
 DB 1124 GTGCTATCTGCCCCCTCCCTGACCCCAAGTACAGATTCCTTACAGTCCACACAGGT 1183
 QY 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPhe 376
 DB 1184 ACTGAGTGTACTGACTGCTGACAGAGATGACAGATGCCCGGCTCCCAATTTCCAC 1243
 QY 377 LeuAsnTyrrArgIleValGluGlnThrProLysLeuProMetAspGly---LeuPheLeu 395
 DB 1244 GTTGTATATGATCTCTGAGCCCTGAGCCCTGAGATGAGGTATAGGGAGAGCCCTTCAG 1303
 QY 396 IleGlnThrTyrrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys---GlnAsp 414
 DB 1304 GTGAGCCCACTTCAGAGGTGACGCTGGGGGTCTCCCATCTCCAGAGCCACAGAC 1363
 QY 415 ThrProGlnTyrrAsnLeuThrIleGluValSerAspLysAsp-----PheLysThrLeu 432
 DB 1364 ATCTGCTTGTGAGTCTGCGCATGACCTGCGAGGCGAGAGGGTGGCTTCACAGACAGC 1423
 QY 433 CysPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
 DB 1424 TGTGAAGTCCAAAGTCGACAGTCCACAGATATCAATGATCACGCCCTGATCATCACTTC 1483
 QY 453 AspTyrrGlyAsnLeuThrIleValAlaAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1484 CAGATTGGGCTTAAAGCTTCTCTGAGATGTGGAGCCGGAGCTCTGCTGGCCATCTTA 1543
 QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrrHis 490
 DB 1544 ACAGCATATGATGTGATCCGAGGCC-----GCCCTCGGCTCATGATTTTGGC 1594
 QY 491 IleIleLysGlyAspSerGluGlyArgLeuGlyValAspThrAspProHisThrAsnThr 510
 DB 1595 ATTGAGAGGGAGACACAGAAAGGAACTTTTGGCTGATTTGGAGCA-----GACTCT 1648
 QY 511 GlyTyrrValIleLys-----LysProLeuAspPheGlnThrAlaAlaValAsnAsn 528
 DB 1649 GGGATGTTGACTCAGACTCTCCAGAACTTATAGGAGAGTCCCAAGTATGAG 1708
 QY 529 IleValPheLysAlaGluAsnProGluProLeuValPheGlyValLysTyrrAsnAlaSer 548
 DB 1709 GTGTTGTGTGTGTGACAGATGTGGGCAACTGTG---GGGCAAGGCCCAAGGCGCTTGA 1765
 QY 549 SerPheAlaLysPheThrIleuIleValThrAspValAsnGluAlaProGlnPheSerGln 568
 DB 1766 GCCACCGCCACGGTGACTGTGATGTGAGAGAGATGTCCACCCCAAGTTGAGACAG 1825

XX 02-APR-2001 (first entry)
 XX DNA encoding protein of the invention #66.
 DE Secreted; transmembrane; gene therapy; 88.
 XX Unidentified.
 OS WO200078961-A1.
 XX 28-DEC-2000.
 XX 18-FEB-2000; 2000WO-US04342.
 XX 23-JUN-1999; 99US-0141037.
 XX 20-JUL-1999; 99US-0144758.
 XX 26-JUL-1999; 99US-0145698.
 XX 01-SEP-1999; 99WO-US20111.
 XX 29-OCT-1999; 99US-0162506.
 XX 30-NOV-1999; 99WO-US28313.
 XX 02-DEC-1999; 99WO-US28551.
 XX 16-DEC-1999; 99WO-US30095.
 XX 05-JAN-2000; 2000WO-US00219.
 XX 06-JAN-2000; 2000WO-US00376.
 XX (GENTH) GENENTECH INC.
 XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Feng S;
 PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillman KJ,
 PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;
 PI Watanabe CK, Williams PM, Wood WI;
 DR WPI; 2001-071395/08.
 XX Secreted and transmembrane proteins and nucleic acids designated PRO,
 PT useful as hybridization probes, in chromosome and gene mapping and gene
 XX therapy -
 XX Claim 2; Fig 131; 787pp; English.
 XX The present invention relates to secreted and transmembrane proteins.
 CC These proteins and the DNA encoding them may be used as hybridization
 CC probes, in chromosome and gene mapping and in the generation of
 CC anti-sense RNA and DNA. They may also be used to generate either
 CC transgenic animals or knockout animals which are in turn useful for
 CC development and screening of therapeutically useful reagents.
 CC The nucleic acids may also be used in gene therapy.
 CC
 XX Sequence 2848 BP; 607 A; 873 C; 828 G; 540 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 2.03e-69 Length: 2848
 Score: 908.50 Matches: 237
 Percent Similarity: 50.51% Conservative: 159
 Best Local Similarity: 30.23% Mismatches: 335
 Query Match: 21.03% Indels: 53
 DB: Gaps: 21
 US-10-025-380-1081 (1-832) x AAF54364 (1-2848)
 QY 62 LeuThrGlyGlu-----ThrAspAsnIlePheValIleGluArgGlu---Gly 76
 DB 296 CTCTCAGGAGCTCAGGAGCAAGCACTGAGGCGCCATTCTCTATGATTCACAGATTCTGCG 355
 QY 77 LeuLeuTyTyrAsnArgAlaLeuAspArgGluThrArgSerThrHisAsnLeuGlnVal 96
 DB 356 TTCTCTCTGCTGACCGAGGCGCTGAGCCGAGAGAGAGAGAGATACAGTACCAAGTCTC 415
 QY 97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
 DB 416 ACCCTGAGATGAGATGAGATGATCTGTGTGGGGTTCACAGCCTGTGCTTGTGACGTG 475

QY 117 LysAspIleAsnAspAsnArgProThrPheLeuGlnSerLysTyrgluGlySerValArg 136
 DB 476 AAGATGAGAAATATACCGAGGTGCGCCATTCTCTCAAGCCATCTACAGAGTCTGGCTGAGC 535
 QY 137 GlnAsnSerArgProGlyLysProPheLeuTyTyrValAsnAlaThrAspLeuAspPro 156
 DB 536 CCGGTGACCGAGGCGCTGACATCCCTTCTCTTCTGAGGCTTCAGACCGGATGAGCCA 595
 QY 157 AlaThrProAsnGlyLeuTyTyrGlnIleValIleGlnLeuProMetIleAsnAsn 176
 DB 596 GGCACAGCCACTCGGATCTTCGATTCACATCTCTGAGCTCAGGCTCCAGCCAGCTTCC 655
 QY 177 ValMetTyPheGlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGluGlySer 196
 DB 656 CCAGACATGTTCCAGCTGGAGCGCTGGGCGCTCTGGCCCTCAGCCCAAGGGAGG 715
 QY 197 GlnGluLeuAsnProAlaLysAsnProSerTyTyrAsnLeuValIleSerValLysAspMet 216
 DB 716 ACCAGCCTTGACCCAGCCCTGGAGAGAGACTTACAGCTGTGTACAGTCAAGACATG 775
 QY 217 GlyGlyGlnSerGluAsnSerPheSerAspThrThrSerValAspIleIleValThrGlu 236
 DB 776 GGTGACCAAGGCGC--TCAGGCCACAGGCCACTGCCACCGTGAAGCTTCATCATAGAG 832
 QY 237 AsnIleTyPheValAlaProLysProValAlaGluMetValGluAsnSerThrAspProHisPro 256
 DB 833 AGCACCTGGGTGTCCTTACAGCTTACACCTGCGAGAGATCTCAAGCTCATATACCG 892
 QY 257 IleLysIleThrGlnValArgTrpAsnAspProGlyAlaGlnTySerLeuValAspLys 276
 DB 893 CACCAATGAGCCAGGTATACCTGAGTGGGGGATGATGCATCATACCTG----- 943
 QY 277 GluLysLeuProArgPheProPheSerIleAspGlnGluGlyAspIleTyTyrValThrGln 296
 DB 944 GAGAGCCATCCCGGGAGCCCTTGAAGTAAATGACAGAGGAAACCTTACGTGACCAAC 1003
 QY 297 ProLeuAspArgGlnGluLysAspAlaTyTyrValPheTyTyrAlaValAlaLysAspGluTyTyr 316
 DB 1004 GAGCTGAGAGAGAAACCCAGGCTGAGTACCTGCTCCAGGTGGGCTCAGATTCCCAT 1063
 QY 317 GlyLysProLeuSerTyTyrProLeuGlnIleHisValIleValLysAspIleAsnAspAsn 336
 DB 1064 GGGAGAGACTATCGGCGCCCTGAGCTGAGCTGAGTGTGATGAGATGACCAAC 1123
 QY 337 ProProThrCyPheProSerProValThrValPheGluValGlnGluAsnGluGluGly 356
 DB 1124 GTGCTATCTGCTCCCTCCGTCAGCCACAGTCAAGATCCAGTCAAGTCCACAGGT 1183
 QY 357 AsnSerIleGlyThrLeuThrAlaHisAspArgAspGluGluAsnThrAlaAsnSerPhe 376
 DB 1184 ACTGAAGTACATGACTGTACAGAGAGATGACAGATGCCCGGCTCCCAATTCCAC 1243
 QY 377 LeuAsnTyTyrArgIleValGluGlnThrProLysLeuProMetAspGly---LeuPheLeu 395
 DB 1244 GTGTGTATACAGCTCTGAGCCCTGAGCTGAGGATGAGGATGAGGAGAGAGCTTCCAG 1303
 QY 396 IleGlnThrTyTyrAlaGlyMetLeuGlnLeuAlaLysGlnSerLeuLys---GlnAsp 414
 DB 1304 GTGAGCCCACTTCAAGAGGTGAGCGTGGGGGTCTCCACTCCAGACAGCCGAGAAC 1363
 QY 415 ThrProGlnTyTyrAsnLeuThrIleGluValSerAspLysAsp-----PheLysThrLeu 432
 DB 1364 ATTCGCTCTGTGTGCTGCGCAATGACCTGGAGGCGAGAGGGTGGCTTCACAGACAG 1423
 QY 433 CyPheValGlnIleAsnValIleAspIleAsnAspGlnIleProIlePheGluLysSer 452
 DB 1424 TGTGAAGTGAAGTGCAGCTCAGATATCATATGATCAAGCCCTGAGTTCACTTCC 1483
 QY 453 AspTyTyrGlyAsnLeuThrLeuAlaGluAspThrAsnIleGlySerThrIleLeuThrIle 472
 DB 1484 CAGATTGGGCTTATTAAGCTCTCCCTGAGATGTGAGGCCGAGGACTTGTGTGATCATGTA 1543
 QY 473 GlnAlaThrAspAlaAsp---GluProPheThrGlySerSerLysIleLeu---TyrHis 490

```

Db      1544  ACAGCCATTGATGTGACCTGAGCC-----GCCTCCGCCCAAGATTGGC 1594
Qy      491  ILeIlelyglYapSerGluGlyArgLeuGlyValAspThrAspProH1srHAsnThr 510
Db      1595  ATTGAGAGGGAGACAGAAAGGAGCTTTGGCTTGATGGAGCC-----GACTCT 1648
Qy      511  GLyTyValIleIleIlely-----LysProLeuAspPheGluThrAlaIleValSerAsn 528
Db      1649  GGGCATTTGACTGACAGACTCTGCAAGAACTCATGATAGGAGCATCCAGTCATGAG 1708
Qy      529  ILeValPheLyAlaGluAsnProGluProLeuValPheGlyValLysTyAsnAlaSer 548
Db      1709  GTGCTGTGTGGTGGTGCAGAGTGTGGCAGAGCTGTG---GGGCGAGGCCAGGCCCTGGA 1765
Qy      549  SerPheAlaLysPheThrLeuIleValThrAspValAsnGluAlaProGluPheSerGln 568
Db      1766  GCCACCCCGCAAGCTGCTGTAGTGAGAGAGCTGATGCCACCCCAAGTTGGACCAAG 1825
Qy      569  HisValPheGlnAlaLysValSerGluAspValAlaIleGlyThrLysValGlyAsnVal 588
Db      1826  GAGAGCTACGAGGCGCTGTCCTCATGATGCCACGCCGCTCTTCTGCTGACCATC 1885
Qy      589  ThrAlaLysAspProGluGlyLeuAspIleSerTySerLeuArgGlyAspThrArgGly 608
Db      1886  CAGCCCTCCGACCCCATCAGCCGAAACCTCAGTTCCTCCATGATGATCAGAGGCG 1945
Qy      609  TrpLeuLysIleAsnH1sValThrGlyGluIlePheSerValAlaProLeuAsp--Arg 627
Db      1946  TGGCTCTGCATTGAGAAATTCTCCGGGAGGTGCACACGCCCATGCTGACAGAGGCGCC 2005
Qy      628  GluAlaGlySerProTyArgValGluValAlaThrGluValGlyGlySerSerLeu 647
Db      2006  CAGCTGAGGACACTTACAGCTGCTGTGAGGCGCAGGATACAGCC----- 2053
Qy      648  SerSerValSerGluPheH1sLeuIleLeuMetAspValAsnAspAsnProArgLeu 667
Db      2054  -----CTGACTCTTCCCTGCGCTCCCA----- 2080
Qy      668  AlaLysAspTyThrGlyLeuPhePheCysH1sProLeuSerAlaProGlySerLeuIle 687
Db      2081  -----TACCTGTGCACACCCCGCCCAAGACATGAGCTTATGCTG 2119
Qy      688  Phe---GluAlaThrAspAspAspGlnH1sLeuPheArgGlyProH1sPheThrPheSer 706
Db      2120  AGTGAACCCAGCAAGAACCCGATCTGGCCAGTGGGACGCTCC---TACAGCTTAC 2176
Qy      707  LeuGly---SerGlySerLeuGlnAsnAspTrpGluValSerLysIleAsnGlyThrH1s 725
Db      2177  CTGTGTCACCAACCCACGCGTCAACGGGATTTGGCGCTCCAGACTCTCAATGTTCCAT 2236
Qy      726  AlaArgLeuSerThrArgH1sThrAspPheGluGluArgAlaTyValValLeuIleArg 745
Db      2237  GCCTACCTCAGCTTGGCCCTTCATTTGGGTGAGCAGCAGCTGAACATATCCCGTGTG 2296
Qy      746  IleAsnAspGlyArgProProLeuGluGlyIleValSerLeuProValThrPheCys 765
Db      2297  GTCTGCCCAATGCCCA-----ATGTGGCAGCTCTGGTTCGAGTATGCTGTGT 2347
Qy      766  SerCys---ValGluGlySerCysPheArgProAlaGlyH1sGlnThrGlyIleProThr 784
Db      2348  CGCTGCAACGTGAGGGGCGCATGTCGCAAGGTGGGCCCAAGAGGGGATGCCACG 2407
Qy      785  ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAla 804
Db      2408  AAGCTGTGGCAGTGGGATCTTTGTAGGCACTCTGTAGCAATAGGATCTTCTCATC 2467
Qy      805  ValValPheIleArgIleLysLysAspLysArgLysAspAsnValGluSerAlaGlnAla 824
Db      2468  CTCATTTTCCACCACTGGACCATGTCAAGAAAGGACCCGATCAACAGCAGACAGC 2527
Qy      825  SerGluValLys 828

```

```

Db      2528  GTGCCCCCTGAG 2539
RESULT 14
ABK33629
ID      ABK33629 standard; cDNA, 2848 BP.
XX
AC      ABK33629;
XX
DT      08-MAY-2002 (first entry)
DE      cDNA encoding human PRO protein. Seq ID No 187.
XX
KW      Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW      breast cancer; prostate tumour; rectal tumour; liver tumour;
KW      pericyte cell proliferation; chondrocyte cell proliferation;
KW      tumour necrosis factor-alpha; gene; ss.
XX
OS      Homo sapiens.
XX
PN      WO200208288-A2.
XX
PD      31-JAN-2002.
XX
PF      29-JUN-2001; 2001WO-US21066.
XX
PR      20-JUL-2000; 2000US-219556P.
PR      25-JUL-2000; 2000US-220585P.
PR      25-JUL-2000; 2000US-220605P.
PR      25-JUL-2000; 2000US-220607P.
PR      25-JUL-2000; 2000US-220624P.
PR      25-JUL-2000; 2000US-220638P.
PR      25-JUL-2000; 2000US-220664P.
PR      25-JUL-2000; 2000US-220666P.
PR      25-JUL-2000; 2000US-220893P.
PR      26-JUL-2000; 2000WO-US20710.
PR      28-AUG-2000; 2000WO-US23522.
PR      24-AUG-2000; 2000WO-US23328.
PR      15-SEP-2000; 2000US-000000P.
PR      10-NOV-2000; 2000WO-US30873.
PR      28-NOV-2000; 2000US-253646P.
PR      01-DEC-2000; 2000WO-US32678.
PR      20-DEC-2000; 2000US-0747259.
PR      20-DEC-2000; 2000WO-US34956.
PR      28-FEB-2001; 2001WO-US06520.
PR      10-MAY-2001; 2001US-0854280.
PR      25-MAY-2001; 2001WO-US17092.
XX
PA      (GBTH ) GENENTECH INC.
XX
PI      Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,
PI      Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR      WPI; 2002-172001/22.
XX
DR      P-PSDB; A0083685.
XX
PT      One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT      useful for treating a PRO related disorder and for diagnosing tumours
PT      such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
PT      tumour or liver tumour -
XX
PS      Claim 2; Figure 187; 359pp; English.
XX
CC      The invention relates to one hundred and twenty two nucleic acids
CC      encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
CC      encode human secreted proteins. The PRO nucleic acids, polypeptides,
CC      agonists and antagonists are useful for treating a PRO related disorder.
CC      The PRO polypeptides are useful for diagnosing tumours, especially lung
CC      cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
CC      liver tumour. The PRO polypeptides are useful for stimulating the
CC      proliferation of, or gene expression, in pericyte cells, for stimulating
CC      the proliferation or differentiation of chondrocyte cells, for
CC      stimulating the release of tumour necrosis factor-alpha from human blood,
CC      for stimulating or inhibiting the proliferation of normal human dermal

```



```

Qy 688 Phe---GluAlaThrAspAspArgLniIleuPheArgGlyProHisPheThrPheSer 706
Db 2120 AGTGAACCCGCAAGAGCCCCGATCTGCGCAAGTGGACAGCTCC--TACAGTTTACC 2176
Qy 707 LeuGly---SerGlySerLeuGlnAsnAspTrpGluValSerIleAsnGlyThrHis 725
Db 2177 CTGTGTCGCAACCCAGCGATGCAACGGGATTTGGCGCTCCAGACTCTCAATGTGTTCCAT 2236
Qy 726 AlaArgLeuSerThrArgHisIleThrAspPheGluGluArgAlaValIleuIleArg 745
Db 2237 GCGTACCTCACTTGGCCCTGATGTGGTGGAGCCAGTAAACATTAATCCCGTGCTG 2296
Qy 746 ILAsnAspGlyArgPProPProLeuGluGlyIleValSerLeuPProValThrPheCys 765
Db 2297 GTGAGCCCAATGCCAG-----ATGTGGAGCTCTCTGGTCAAGTATCGTGT 2347
Qy 766 SerCys---ValGluGlySerCysPheArgPProAlaGlyIleGlnThrGlyIleProThr 784
Db 2348 CGGTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGCCGCAATGAAGGCGATCCACG 2407
Qy 785 ValGlyMetAlaValGlyIleLeuLeuThrThrLeuLeuValIleGlyIleIleLeuAla 804
Db 2408 AAGCTGTGGCAGTGGGATCTCTGTAGGCGACCTGTGTAGCAATAGCAATCTTCTCATC 2467
Qy 805 ValValPheIleArgIleIleValAspIleGlyIleValAspAsnValGluSerAlaGlnAla 824
Db 2468 CTCATTTTCACCCATGCAATGATCTCAAGAGAGAGACCGGATCAACGACGACAGAC 2527
Qy 825 SerGluValIleArg 828
Db 2528 GTGCCCTGAAG 2539

RESULT 15
AAS15602
ID AAS15602 standard; CDNA; 2938 BP.
AC AAS15602;
DT 18-JUN-2002 (first entry)
XX
XX Human CDNA encoding cadherin-2 (CDHN-2).
DE
XX
XX ss; cadherin; CDHN-2; nootropic; neuroprotective; cardiac; human;
KW antiinflammatory; gastric; immunostimulant; cytoprotective; gene-therapy;
KW Alzheimer's disease; Parkinson's disease; multiple sclerosis; cancer;
KW atherosclerosis; angiogenesis; musculoskeletal; ataxia; myotonia;
KW gastroenteric; gastritis; insulin dependent diabetes mellitus;
KW Crohn's disease; inflammatory; asthma; rheumatoid arthritis; lupus.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1.161
FT 5'UTR /*tag= a
FT sig_peptide 162..224
FT /*tag= b
FT /*note= "signal peptide"
FT CDS 162..2654
FT /*tag= c
FT /*product= "CDHN-2"
FT /*note= "This sequence is specifically claimed in
FT claim 1."
FT mat_peptide 225..2651
FT /*tag= d
FT /*note= "Mature CDHN-2"
FT 3'UTR 2655..2938
FT /*tag= e
XX
XX WO200179293-A2.
XX
XX 25-OCT-2001.
XX
XX 18-APR-2001; 2001WO-US12687.

```

```

XX
XX 18-APR-2000; 2000US-198466P.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Curtis RAJ;
PI
XX
XX WPI; 2001-611722/70.
DR
XX
XX P-FSDB; AAU09959.
DR
XX
XX Isolated cadherin polypeptide useful for the treatment of central
PT nervous system, cardiovascular, musculoskeletal, gastrointestinal,
PT inflammatory, immune system and cell proliferation disorders -
XX
XX Claim 1, Figure 7; 147pp; English.
PS
XX
XX This sequence represents the CDHN-2 gene of the invention, this
CC invention comprises the nucleotide and protein sequences of two
CC members of the cadherin family of molecules CDHN-1 and CDHN-2.
CC Also included in the invention are a vector containing the CDHN
CC gene, a host cell transfected with the vector and a method for
CC producing the CDHN proteins by culturing the host cell line. Also
CC included is an antibody which selectively binds to the protein
CC The CDHN-1 protein may have nootropic, neuroprotective, cardiac,
CC antiinflammatory, gastric, immunostimulant and cytoprotective activity
CC and can be used as a cadherin-modulator using antisense-therapy or
CC gene-therapy. These proteins are useful for identifying compounds
CC which bind or modulate CDHN-1. The proteins, nucleic acids and
CC antibodies are useful for treating a subject with a disorder
CC characterised by aberrant or unwanted cadherin protein or nucleic
CC acid activity. These disorders include central nervous system
CC (e.g. Alzheimer's disease, Parkinson's disease, multiple sclerosis),
CC cardiovascular (e.g. atherosclerosis, angiogenesis), musculoskeletal
CC (e.g. ataxia, myotonia) gastrointestinal (e.g. gastritis, insulin
CC dependent diabetes mellitus, Crohn's disease) (e.g. inflammatory and
CC immune system (e.g. asthma, rheumatoid arthritis, lupus) or cell
CC proliferation disorders (e.g. cancers, leukaemia).
XX
XX Sequence 2938, BP; 723 A; 847 C; 766 G; 602 T; 0 other;
SQ
XX
XX Alignment Scores:
Pred. No.: 8,896-67 Length: 2938
Score: 878.50 Matches: 248
Percent Similarity: 47.72% Conservative: 140
Best Local Similarity: 30.50% Mismatches: 358
Query Match: 20.33% Indels: 67
DB: Gaps: 22
US-10-025-380-1081 (1-832) x AAS15602 (1-2938)
Qy 62 LeuThrGlyGluThrAsp-----AsnIlePheValIleGluArgGlu---Gly 76
Db 336 CTATCAGAGACTCTCAACACGCGAGATCAAAACCTTGTCTGGACACAGCTTGGC 335
Qy 77 LeuLeuTyTYrAsnArgAlaLeuAspArgGluThrArgSerThrIleAsnLeuGlnVal 96
Db 396 TTCTCTAGTGGCGACAGAGACCTCGAACCGGAGAGAGAACAAATACCAATACAGGTC 455
Qy 97 AlaAlaLeuAspAlaAsnGlyIleIleValGluGlyProValProIleThrIleGluVal 116
Db 456 ACCTTGAGCTGTAGATGACCTATCTTGTGGGTCACAGCTTGACTGTGCAATG 515
Qy 117 LysAspIleAsnAspAsnArgPProThrPheLeuGlnSerIleTyArgIleGlySerValArg 136
Db 516 AADGATGAGATATACAGATACCCCAATTTCCAGGCACTTACAGAGCTCAGCTGAGC 575
Qy 137 GlnAsnSerArgPProGlyIleYsPProPheLeuTyValAsnAlaThrAspLeuAspPro 156
Db 576 CAGGACACGAGGCTGGGGTCCCTTCTCTTCTTCTGAGGCTCTGATGGGAGTACCA 635
Qy 157 AlaThrProAsnGlyLeuLeuTyTYrGlnIleValIleGlnLeuPProIleIleAsn 176
Db 636 GGCACAGCTAACTCCAGCTTCCCTTCCACATTTGTGAGCAGTCCACCTCAGCCTTTA 695

```

QY 177 ValMetYrPheGlnIleAsnAsnLysThrGlyAlaIleSerLeuThrArgGlnGlySer 196
 DB 696 CCAGACATGTTCCAGCTGAGACCCCTCACCTAGGGGGCTGTGCTTATGCCAGTGAAGC 755
 QY 197 GlnGlnLeuAsnProAlaLysAsnProSerTyrAsnLeuValIleSerValLysAspMet 216
 DB 756 ACCAGCTTACCATGCTTGAAGACCTTACCATGCTTGAAGACCTTGAAGACCTTGAAG 815
 QY 217 GlyGlyLysSerGlnAsnSerPheSerAspThrThrSerValIleLeuIleValThrGln 236
 DB 816 GGTGACCAAGCTTCCAGGCCACAGCTATTGCACT--GTAGAGATCCATATGATGAG 872
 QY 237 AsnIleTyrLysAlaProLysProValGlnMetValGlnAsnSerThrAspProIlePro 256
 DB 873 AACAGCTGGGACCCCTGAGAGCCCTTCACTGGGACAGAAATCCAAAGTTGGTACCA 932
 QY 257 IleLysIleThrGlnValArgTyrAsnAspProGlyAlaGlnTyrSerLeuValAspLys 276
 DB 933 CACAGCATTCGCCAGGTGCACTGAGTGAAGAGACCTGCACTACCACTG----- 983
 QY 277 GlnLysLeuProArgPheProPheSerIleAspGlnGlnLysAspIleTyrValThrGln 296
 DB 984 GAGAGCCAGCTCCAGGACCTTTCGATGTGATACAGAGGGAGATGCTCCATGTTCACATG 1043
 QY 297 ProLeuAspArgGlnGlnLysAspAlaTyrValPheTyrAlaValAlaLysAspGlnTyr 316
 DB 1044 GAGCTGGACCGGAGGAGCGGCTGAGTACCACTCCAACTCCGAGCTCGAAATTCCTCAT 1103
 QY 317 GlyLysProLeuSerTyrProLeuGlnIleIleValLysValLysAspIleAsnAspAsn 336
 DB 1104 GGTGAGACTACACAGAACCCCTGAGTTGCACTGCTGTGTATGATGAACAGCAAT 1163
 QY 337 ProProThrCysProSerProValThrValPheGlnValGlnGlnLysGlnLysGlnGly 356
 DB 1164 GCACCTGTCTGCTCCCAATGACCCCAAGTCAACATCCCTAGCTCAGCCCCCAAGA 1223
 QY 357 AsnSerIleGlyThrLeuThrAlaIleAspArgAspGlnGlnLysAsnThrAlaAsnSerPhe 376
 DB 1224 ACTGAATAGCCAGGCTCTCAGCAGAGATTTGATGCTCCCGGTGACCCCAATTCCTCC 1283
 QY 377 LeuAsnTyrArgIleValGlnGlnThrProLysLeuProMetAspGly--LeuPheLeu 395
 DB 1284 ATTGTATATCATGTTGTGCTGAGCCCTGAGCCCTGAGAGGGGGCTGAACAAAGCTTCGAG 1343
 QY 396 IleGlnThrTyrAlaGlyMetLeuGlnLeu-----AlaLysGlnSer 409
 DB 1344 TTGATCCGACCTCAGGAGCTGTAACATGGGAATGCCCACTCCATGCTGCGCAGAGCT 1403
 QY 410 LeuLysLeuGlnAspThrProGlnTyrAsnLeuThrIleGlnVal-----SerAspLys 427
 DB 1404 ATCTGCTTCAG-----GTCTGCTGCTTTCACCTAGCAGGATCAGAGAGT 1448
 QY 428 AspPheLysThrLeuCysPheValGlnIleAsnValIleAspIleAsnAspGlnIlePro 447
 DB 1449 GGCCTCAGACGACATGTGAGTGAAGCATGTGTCAGACGTCACAAACATATGCCCT 1508
 QY 448 IlePheGlnLysSerAspTyrGlyAsnLeuThrLeuAlaGlnAspThrAsnIleGlySer 467
 DB 1509 GAGTTATCATATCCAGATGGGCTGTAACCTTCTGAGGATGTAAACCTGGAGCT 1568
 QY 468 ThrIleLeuThrIleGlnAlaThrAspAlaAsp--GluProPheThrGlySerLys 486
 DB 1569 CTGGTGGCAACTCATATGCGCACTGATGCTGACCTTGAACT-----GCTTCCTCC 1619
 QY 487 IleLeu--TyrHisIleIleLysGlyAspSerGlnGlyArgLeuGlyValAspThrAsp 505
 DB 1620 CTATAGACTTGGCATTTGAAGAAGAGACCCAGAGGATCTTGAACCTGCTGAGAG 1679
 QY 506 ProHisThrAsnThrGlyTyrValIleIleLysLysProLeuAspPheGlnThrAlaAla 525
 DB 1680 CCAAGACTCCGACCATGTCCAGCTCAGATCCGGAAGAACTCAGCTATGAGGAGCTCCT 1739

QY 526 ValSerAsnIleValPheLysAlaGlnAsnProGlnProLeuValPheGlyValLysTyr 545
 DB 1740 GATCACAAGGT 1796
 QY 546 AsnAlaSerSerPheAlaLysPheThrLeuIleValIleThrAspValAsnGlnAlaProGln 565
 DB 1797 GGCCTCGACCCACACACACAGTACTATATGATGAGAGGGGTGTGTGTGTGTGTGTGTGTGT 1856
 QY 566 PheSerGlnHisValPheGlnAlaLysValSerGlnAspValAlaIleGlyThrLysVal 585
 DB 1857 TTGACACAGAGAGCTATGAGACACACATCCAGTACAGACCCCACTGAGCTCTCTCTG 1916
 QY 586 GlyAsnValThrAlaLysAspProGlnGlyLeuAspIleSerTyrSerLeuAspGlyAsp 605
 DB 1917 CTGACCATCAGCCCTCAGACCCCATGAGAGAACCTTCAGTTCCTCGTGCATATGAC 1976
 QY 606 ThrArgGlyTyrPheLysIleAspHisValIleThrGlyGlnIlePheSerValAlaProLeu 625
 DB 1977 TCAGAGGGCTGCTGTATCAAGAGGTCTGTGGAGGTATACACAGCCCACTGCTCTG 2036
 QY 626 Asp--ArgGlnAlaGlySerProTyrArgValGlnValAla-----ThrGlnVal 642
 DB 2037 CAGGCTGCCAGCTTGAACACATACACATACAGTCTTGTGAGGCCCAAGACACAGATTAAG 2096
 QY 643 GlyLysSerSerLeuSerSerValSerGlnPheIleLeu----- 656
 DB 2097 CCAAGACTAGACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2156
 QY 657 LeuMetAspValAsnAspAsnProProArgLeuAlaLysAspTyrThrGlyLeuPhePhe 676
 DB 2157 GCATTGACTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2192
 QY 677 CysHisProLeuSerLysAlaProGlySerLeuIlePheGlnAlaThrAspAspAspGlnHis 696
 DB 2193 TGTACACCCCGCCAGACATACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2252
 QY 697 LeuPheArgGlyProHisPheThrPheSerLeuGly--SerGlySerLeuGlnAsnAsp 715
 DB 2253 GCCAAGAGAAATGTGTCTTCAACAGCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2312
 QY 716 TrpGlnValSerLysIleAsnGlyThrHisAlaArgLeuSerThrArgHisThrAspPhe 735
 DB 2313 TGGCGCTCAGACCTTCAACAGTTCACAGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2372
 QY 736 GlnGlnArgAlaTyrValValLeuIleArgIle--AsnAspGlyLysArgProProLeu 754
 DB 2373 GAGCTGTGAAATACATGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2432
 QY 755 GlnGlyIleValSerLeuProValThrPheCysSerCys--ValGlnGlySerCysPhe 773
 DB 2433 CAG-----GTCAAAGT 2480
 QY 774 ArgProAlaGlyHisGlnThrGlyIleProThrValGlyMetAlaValGlyIleLeuLeu 793
 DB 2481 CCCAAGT 2540
 QY 794 ThrThrLeuLeuValIleGlyIleIleLeuAlaValAlaPhe-----Ile 808
 DB 2541 GGCACCTTGGACCGAGTATGCTTTCATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2600
 QY 809 ArgIleLysLysAspLysGlyLysAspAsnVal----- 819
 DB 2601 AGGAAGACCTGTGATCAGCAGACACACAGCTGCTGTGAAGCAGCGGTGTGAATGATC 2660
 QY 820 GluSerAlaGlnAlaSerGlnValLysProLeuArgSer 832
 DB 2661 CAAAGACCCCACTGGAGAGTTGGCCCACTGCTCTT 2699

Search completed: June 22, 2003, 01:20:05
 Job time : 546 secs

THIS PAGE BLANK (USPS)

BEST AVAILABLE COPY